IV

INTERNAL ORGANIZATION OF THE CELL

Part IV
Chapters

10 Membrane Structure

11 Membrane Transport of Small Molecules and the Electrical Properties of Membranes

12 Intracellular Compartments and Protein Sorting

13 Intracellular Vesicular Traffic

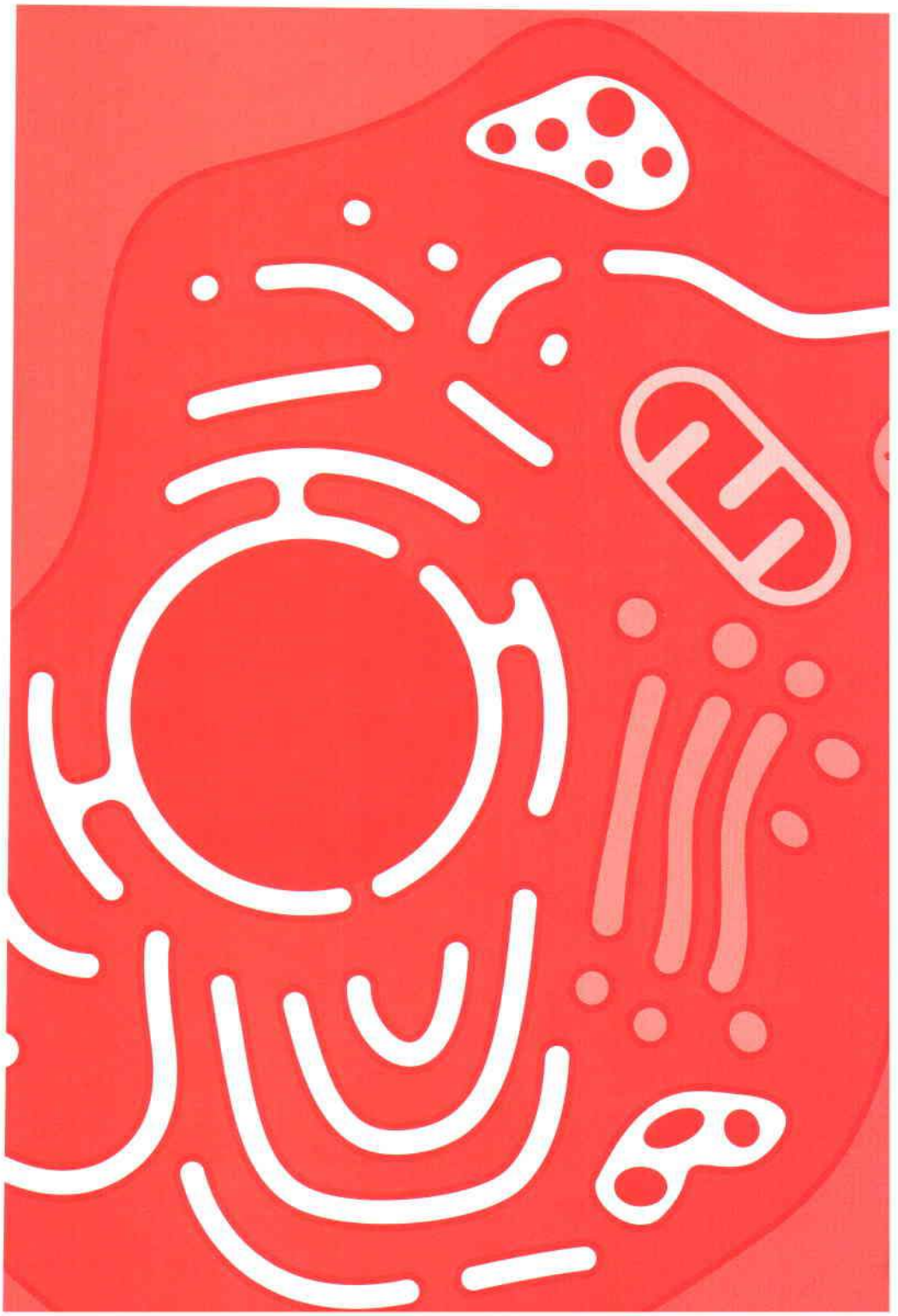
14 Energy Conversion: Mitochondria and Chloroplasts

15 Mechanisms of Cell Communication

16 The Cytoskeleton

17 The Cell Cycle

18 Apoptosis



10

Membrane Structure

In This Chapter

THE LIPID BILAYER 617

MEMBRANE PROTEINS 629

Cell membranes are crucial to the life of the cell. The **plasma membrane** encloses the cell, defines its boundaries, and maintains the essential differences between the cytosol and the extracellular environment. Inside eucaryotic cells, the membranes of the endoplasmic reticulum, Golgi apparatus, mitochondria, and other membrane-enclosed organelles maintain the characteristic differences between the contents of each organelle and the cytosol. Ion gradients across membranes, established by the activities of specialized membrane proteins, can be used to synthesize ATP, to drive the transmembrane movement of selected solutes, or, as in nerve and muscle cells, to produce and transmit electrical signals. In all cells, the plasma membrane also contains proteins that act as sensors of external signals, allowing the cell to change its behavior in response to environmental cues, including signals from other cells; these protein sensors, or *receptors*, transfer information—rather than molecules—across the membrane.

Despite their differing functions, all biological membranes have a common general structure: each is a very thin film of lipid (fatty) and protein molecules, held together mainly by noncovalent interactions (**Figure 10–1**). Cell membranes are dynamic, fluid structures, and most of their molecules move about in the plane of the membrane. The lipid molecules are arranged as a continuous double layer about 5 nm thick. This *lipid bilayer* provides the basic fluid structure of the membrane and serves as a relatively impermeable barrier to the passage of most water-soluble molecules. Protein molecules that span the lipid bilayer (*transmembrane proteins*; see Figure 10–1) mediate nearly all of the other functions of the membrane, transporting specific molecules across it, for example, or catalyzing membrane-associated reactions such as ATP synthesis. In the plasma membrane, some transmembrane proteins serve as structural links that connect the cytoskeleton through the lipid bilayer to either the extracellular matrix or an adjacent cell, while others serve as receptors to detect and transduce chemical signals in the cell's environment. As would be expected, it takes many different membrane proteins to enable a cell to function and interact with its environment, and it is estimated that about 30% of the proteins encoded in an animal cell's genome are membrane proteins.

In this chapter, we consider the structure and organization of the two main constituents of biological membranes—the lipids and the proteins. Although we focus mainly on the plasma membrane, most concepts discussed apply to the various internal membranes in cells as well. The functions of cell membranes are considered in later chapters: their role in ATP synthesis, for example, is discussed in Chapter 14; their role in the transmembrane transport of small molecules in Chapter 11; and their roles in cell signaling and cell adhesion in Chapters 15 and 19, respectively. In Chapters 12 and 13, we discuss the internal membranes of the cell and the protein traffic through and between them.

THE LIPID BILAYER

The **lipid bilayer** provides the basic structure for all cell membranes. It is easily seen by electron microscopy, and its structure is attributable exclusively to the

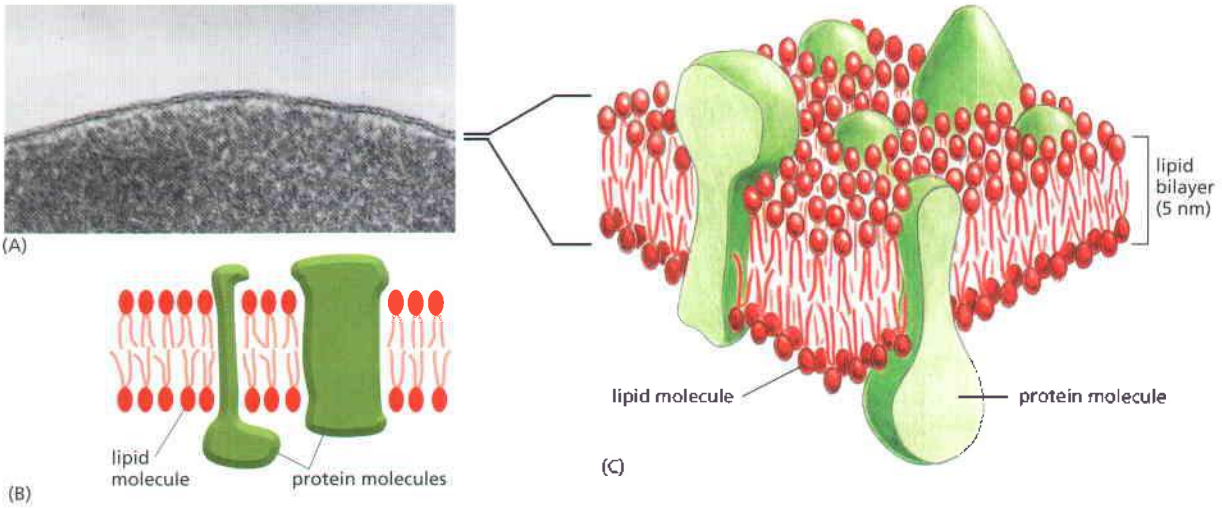


Figure 10–1 **Three views of a cell membrane.** <TAGC> (A) An electron micrograph of a plasma membrane (of a human red blood cell) seen in cross section. (B and C) These drawings show two-dimensional and three-dimensional views of a cell membrane and the general disposition of its lipid and protein constituents. (A, courtesy of Daniel S. Friend.)

special properties of the lipid molecules, which assemble spontaneously into bilayers even under simple artificial conditions.

Phosphoglycerides, Sphingolipids, and Sterols Are the Major Lipids in Cell Membranes

Lipid molecules constitute about 50% of the mass of most animal cell membranes, nearly all of the remainder being protein. There are approximately 5×10^6 lipid molecules in a 1 μm × 1 μm area of lipid bilayer, or about 10^9 lipid molecules in the plasma membrane of a small animal cell. All of the lipid molecules in cell membranes are **amphiphilic**—that is, they have a **hydrophilic** ("water-loving") or *polar* end and a **hydrophobic** ("water-fearing") or *nonpolar* end.

The most abundant membrane lipids are the **phospholipids**. These have a polar head group and two hydrophobic *hydrocarbon tails*. In animal, plant, and bacterial cells, the tails are usually fatty acids, and they can differ in length (they normally contain between 14 and 24 carbon atoms). One tail typically has one or more *cis*-double bonds (that is, it is *unsaturated*), while the other tail does not (that is, it is *saturated*). As shown in **Figure 10–2**, each *cis*-double bond creates a small kink in the tail. Differences in the length and saturation of the fatty acid tails influence how phospholipid molecules pack against one another, thereby affecting the fluidity of the membrane, as we discuss later.

The main phospholipids in most animal cell membranes are the **phosphoglycerides**, which have a three-carbon *glycerol* backbone (see Figure 10–2). Two long-chain fatty acids are linked through ester bonds to adjacent carbon atoms of the glycerol, and the third carbon atom is attached to a phosphate group, which in turn is linked to one of several different types of head group. By combining several different fatty acids and head groups, cells make many different phosphoglycerides. *Phosphatidylethanolamine*, *phosphatidylserine*, and *phosphatidylcholine* are the main ones in mammalian cell membranes (**Figure 10–3A–C**)

Another important phospholipid, called *sphingomyelin*, is built from *sphingosine* rather than glycerol (Figure 10–3D–E). Sphingosine is a long acyl chain with an amino group (NH_2) and two hydroxyl groups (OH) at one end of the molecule. In sphingomyelin, a fatty acid tail is attached to the amino group, and a phosphocholine group is attached to the terminal hydroxyl group, leaving one hydroxyl group free. The free hydroxyl group contributes to the polar properties of the adjacent head group, as it can form hydrogen bonds with the head group of a neighboring lipid, with a water molecule, or with a membrane protein. Together, the phospholipids phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, and sphingomyelin constitute more than half the mass of lipid in most mammalian cell membranes (see Table 10–1).

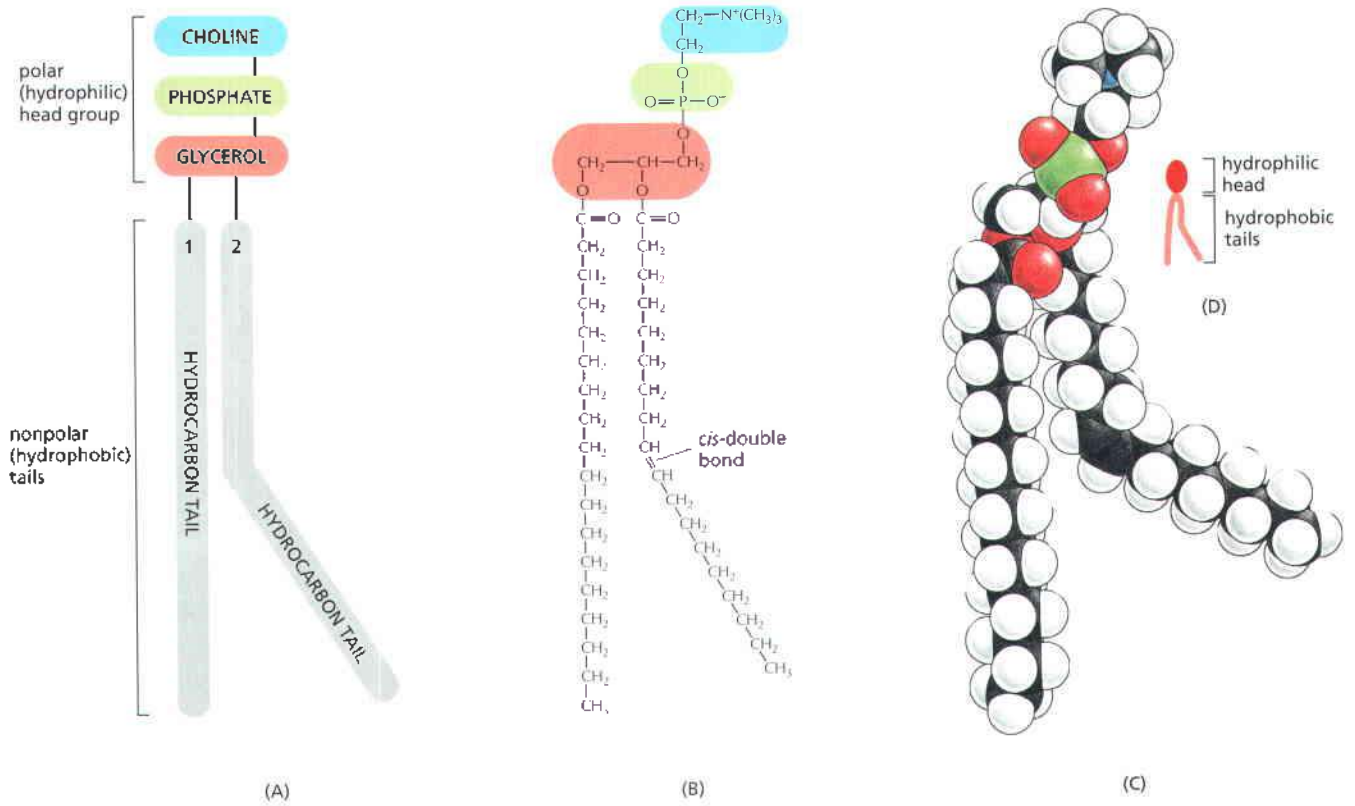


Figure 10–2 The parts of a phosphoglyceride molecule. This example is a phosphatidylcholine, represented (A) schematically, (B) by a formula, (C) as a space-filling model, and (D) as a symbol. The kink resulting from the *cis*-double bond is exaggerated for emphasis.

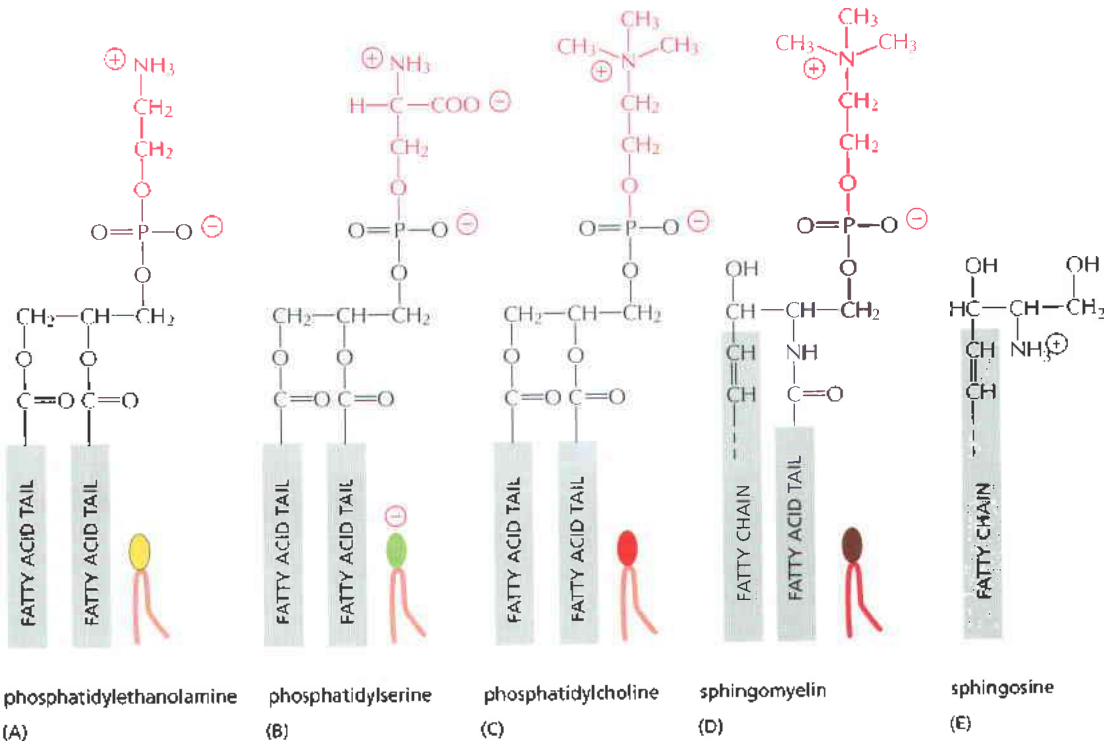


Figure 10–3 Four major phospholipids in mammalian plasma membranes. Different head groups are represented by different colors in the symbols. The lipid molecules shown in (A–C) are phosphoglycerides, which are derived from glycerol. The molecule in (D) is sphingomyelin, which is derived from sphingosine (E) and is therefore a sphingolipid. Note that only phosphatidylserine carries a net negative charge, the importance of which we discuss later; the other three are electrically neutral at physiological pH, carrying one positive and one negative charge.

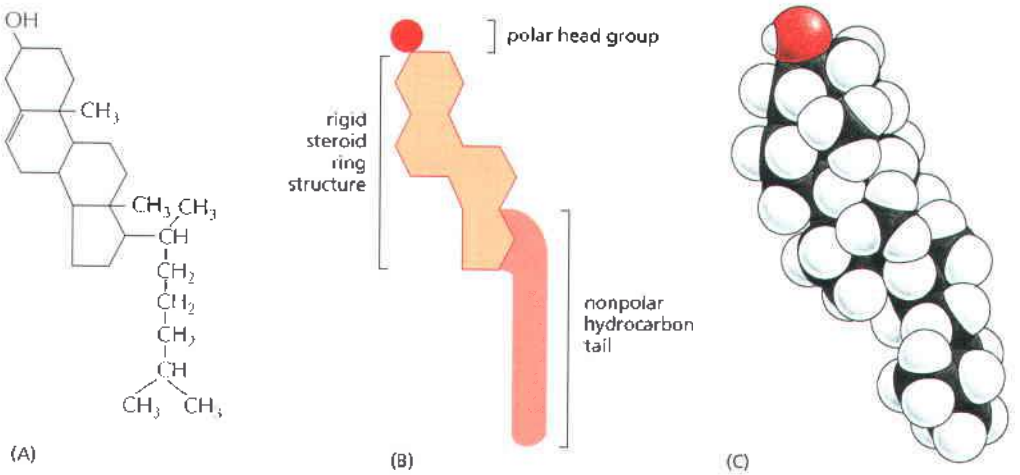


Figure 10–4 **The structure of cholesterol.** Cholesterol is represented (A) by a formula, (B) by a schematic drawing, and (C) as a space-filling model.

In addition to phospholipids, the lipid bilayers in many cell membranes contain *cholesterol* and *glycolipids*. Eucaryotic plasma membranes contain especially large amounts of **cholesterol** (**Figure 10–4**)—up to one molecule for every phospholipid molecule. Cholesterol is a sterol. It contains a rigid ring structure, to which is attached a single polar hydroxyl group and a short nonpolar hydrocarbon chain. The cholesterol molecules orient themselves in the bilayer with their hydroxyl group close to the polar head groups of adjacent phospholipid molecules (**Figure 10–5**).

Phospholipids Spontaneously Form Bilayers

The shape and amphiphilic nature of the phospholipid molecules cause them to form bilayers spontaneously in aqueous environments. As discussed in Chapter 2, hydrophilic molecules dissolve readily in water because they contain charged groups or uncharged polar groups that can form either favorable electrostatic interactions or hydrogen bonds with water molecules. Hydrophobic molecules, by contrast, are insoluble in water because all, or almost all, of their atoms are uncharged and nonpolar and therefore cannot form energetically favorable interactions with water molecules. If dispersed in water, they force the adjacent water molecules to reorganize into ice-like cages that surround the hydrophobic molecule (**Figure 10–6**). Because these cage structures are more ordered than the surrounding water, their formation increases the free energy. This free energy cost is minimized, however, if the hydrophobic molecules (or the hydrophobic portions of amphiphilic molecules) cluster together so that the smallest number of water molecules is affected.

The hydrophobic and hydrophilic regions of lipid molecules behave in the same way. Thus, lipid molecules spontaneously aggregate to bury their hydrophobic hydrocarbon tails in the interior and expose their hydrophilic heads to water. Depending on their shape, they can do this in either of two ways: they can form spherical *micelles*, with the tails inward, or they can form double-layered sheets, or *bilayers*, with the hydrophobic tails sandwiched between the hydrophilic head groups (**Figure 10–7**).

Being cylindrical, phospholipid molecules spontaneously form bilayers in aqueous environments. In this energetically most favorable arrangement, the hydrophilic heads face the water at each surface of the bilayer, and the hydrophobic tails are shielded from the water in the interior. The same forces that drive phospholipids to form bilayers also provide a self-healing property. A small tear in the bilayer creates a free edge with water; because this is energetically unfavorable, the lipids tend to rearrange spontaneously to eliminate the free edge. (In eucaryotic plasma membranes, the fusion of intracellular vesicles repairs larger tears.) The prohibition of free edges has a profound consequence: the only way for a bilayer to avoid having edges is by closing in on itself and forming a sealed compartment (**Figure 10–8**). This remarkable behavior, fundamental

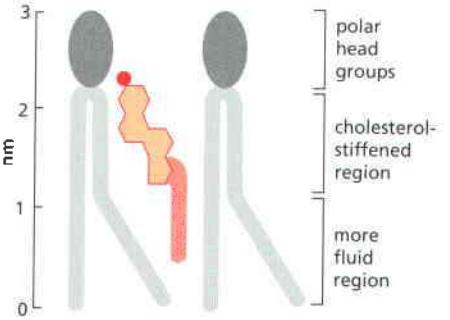


Figure 10–5 **Cholesterol in a lipid bilayer.** Schematic drawing of a cholesterol molecule interacting with two phospholipid molecules in one monolayer of a lipid bilayer.

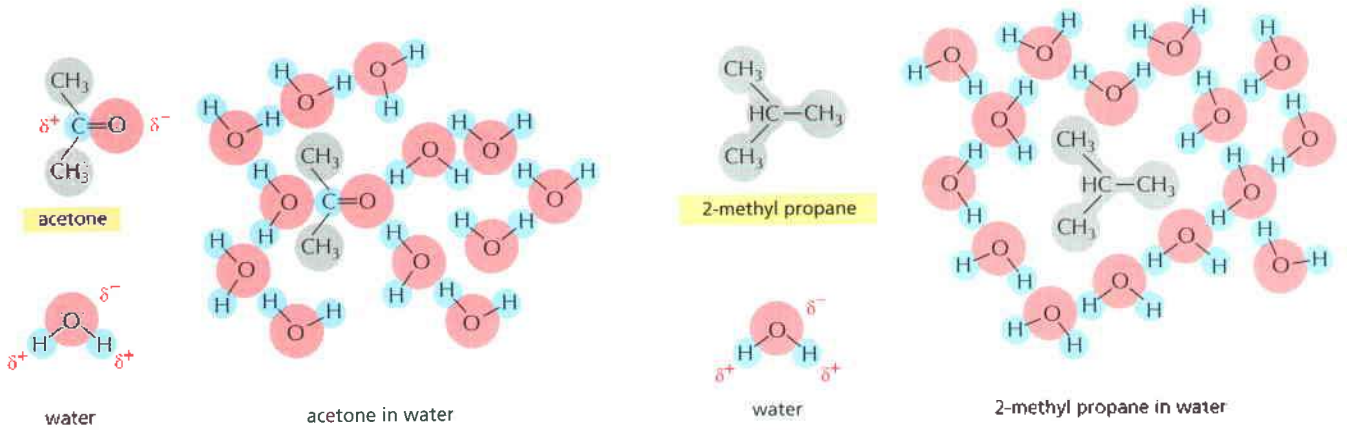


Figure 10–6 How hydrophilic and hydrophobic molecules interact differently with water. (A) Because acetone is polar, it can form favorable electrostatic interactions with water molecules, which are also polar. Thus, acetone readily dissolves in water. (B) By contrast, 2-methyl propane is entirely hydrophobic. Because it cannot form favorable interactions with water, it would force adjacent water molecules to reorganize into icelike cage structures, which increases the free energy. This compound is therefore virtually insoluble in water. The symbol δ^- indicates a partial negative charge, and δ^+ indicates a partial positive charge. Polar atoms are shown in color and nonpolar groups are shown in *gray*.

to the creation of a living cell, follows directly from the shape and amphiphilic nature of the phospholipid molecule.

A lipid bilayer also has other characteristics that make it an ideal structure for cell membranes. One of the most important of these is its fluidity, which is crucial to many membrane functions.

The Lipid Bilayer Is a Two-dimensional Fluid

Around 1970, researchers first recognized that individual lipid molecules are able to diffuse freely within lipid bilayers. The initial demonstration came from studies of synthetic lipid bilayers. Two types of preparations have been very useful in such studies: (1) bilayers made in the form of spherical vesicles, called **liposomes**, which can vary in size from about 25 nm to 1 μm in diameter

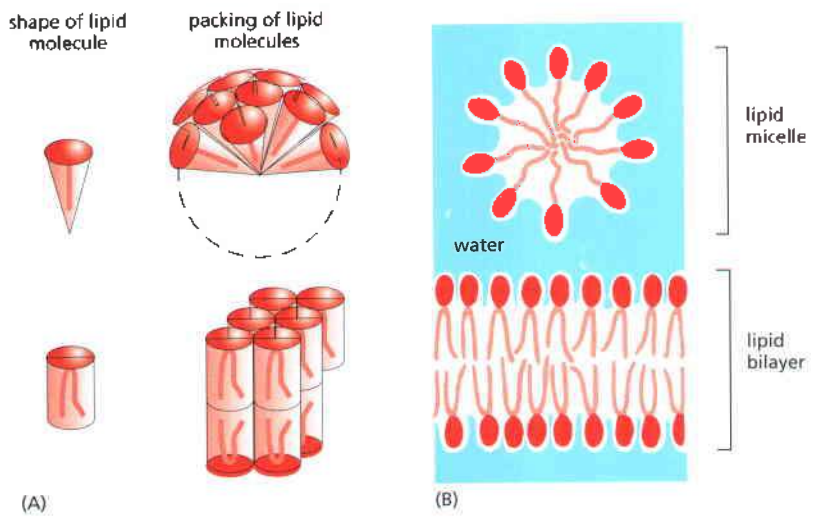


Figure 10–7 Packing arrangements of lipid molecules in an aqueous environment. (A) Cone-shaped lipid molecules *(above)* form micelles, whereas cylinder-shaped phospholipid molecules *(below)* form bilayers. (B) A lipid micelle and a lipid bilayer seen in cross section. Lipid molecules spontaneously form one or the other structure in water, depending on their shape.

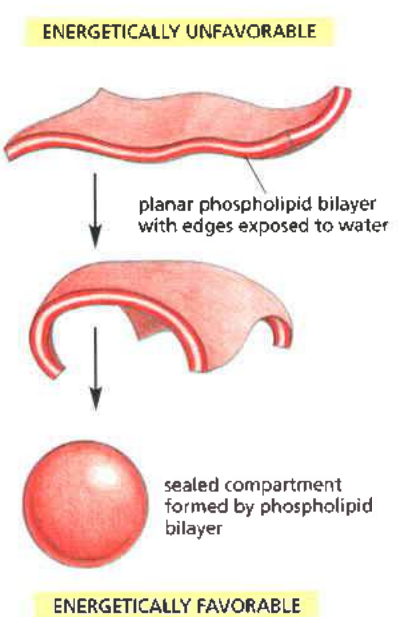


Figure 10–8 The spontaneous closure of a phospholipid bilayer to form a sealed compartment. The closed structure is stable because it avoids the exposure of the hydrophobic hydrocarbon tails to water, which would be energetically unfavorable.

depending on how they are produced (**Figure 10–9**); and (2) planar bilayers, called **black membranes**, formed across a hole in a partition between two aqueous compartments (**Figure 10–10**).

Various techniques have been used to measure the motion of individual lipid molecules and their components. One can construct a lipid molecule, for example, with a fluorescent dye or a small gold particle attached to its polar head group and follow the diffusion of even individual molecules in a membrane. Alternatively, one can modify a lipid head group to carry a "spin label," such as a nitroxyl group (>N–O); this contains an unpaired electron whose spin creates a paramagnetic signal that can be detected by electron spin resonance (ESR) spectroscopy. (The principles of this technique are similar to those of nuclear magnetic resonance, discussed in Chapter 8.) The motion and orientation of a spin-labeled lipid in a bilayer can be deduced from the ESR spectrum. Such studies show that phospholipid molecules in synthetic bilayers very rarely migrate from the monolayer (also called a *leaflet*) on one side to that on the other. This process, known as "flip-flop," occurs less than once a month for any individual molecule, although cholesterol is an exception to this rule and can flip-flop rapidly. In contrast, lipid molecules readily exchange places with their neighbors *within* a monolayer ($\sim 10^7$ times per second). This gives rise to a rapid lateral diffusion, with a diffusion coefficient (D) of about 10^{-8} cm^2/sec, which means that an average lipid molecule diffuses the length of a large bacterial cell (~2 μm) in about 1 second. These studies have also shown that individual lipid molecules rotate very rapidly about their long axis and have flexible hydrocarbon chains. Computer simulations show that lipid molecules in membranes are very disordered, presenting an irregular surface of variously spaced and oriented head groups to the water phase on either side of the bilayer (**Figure 10–11**).

Similar mobility studies on labeled lipid molecules in isolated biological membranes and in living cells give results similar to those in synthetic bilayers. They demonstrate that the lipid component of a biological membrane is a two-dimensional liquid in which the constituent molecules are free to move laterally. As in synthetic bilayers, individual phospholipid molecules are normally confined to their own monolayer. This confinement creates a problem for their synthesis. Phospholipid molecules are manufactured in only one monolayer of a membrane, mainly in the cytosolic monolayer of the endoplasmic reticulum membrane. If none of these newly made molecules could migrate reasonably promptly to the noncytosolic monolayer, new lipid bilayer could not be made. The problem is solved by a special class of transmembrane enzymes called *phospholipid translocators*, which catalyze the rapid flip-flop of phospholipids from one monolayer to the other, as discussed in Chapter 12.

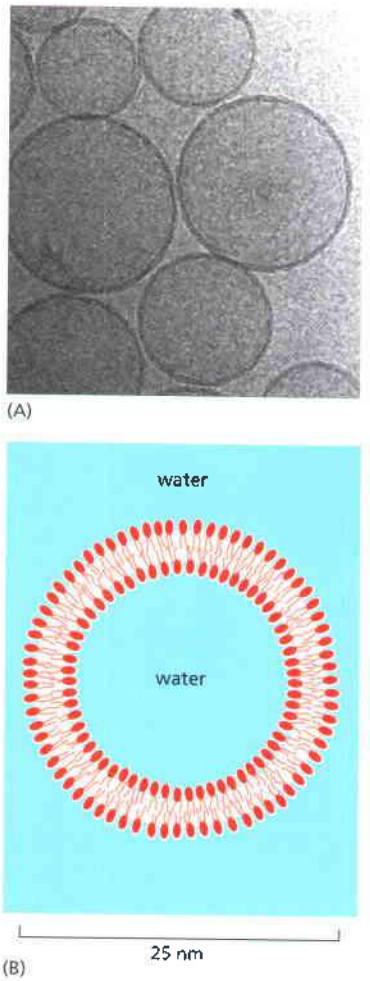


Figure 10–9 **Liposomes.** (A) An electron micrograph of unfixed, unstained phospholipid vesicles—liposomes—in water rapidly frozen to liquid nitrogen temperature. (B) A drawing of a small spherical liposome seen in cross section. Liposomes are commonly used as model membranes in experimental studies. (A, from P. Frederik and W. Hubert, *Meth. Enzymol.* 391:431, 2005. With permission from Elsevier.)

The Fluidity of a Lipid Bilayer Depends on Its Composition

The fluidity of cell membranes has to be precisely regulated. Certain membrane transport processes and enzyme activities, for example, cease when the bilayer viscosity is experimentally increased beyond a threshold level.

The fluidity of a lipid bilayer depends on both its composition and its temperature, as is readily demonstrated in studies of synthetic bilayers. A synthetic bilayer made from a single type of phospholipid changes from a liquid state to a

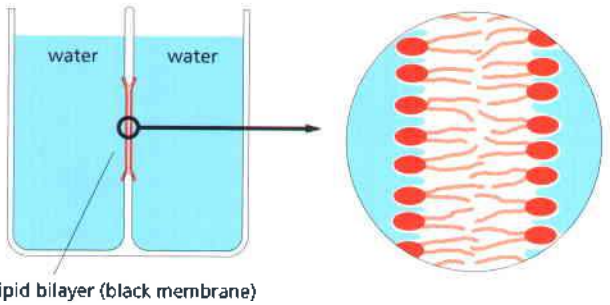


Figure 10–10 **A cross-sectional view of a black membrane, a synthetic lipid bilayer.** This planar bilayer appears black when it forms across a small hole in a partition separating two aqueous compartments. Black membranes are used to measure the permeability properties of synthetic membranes.

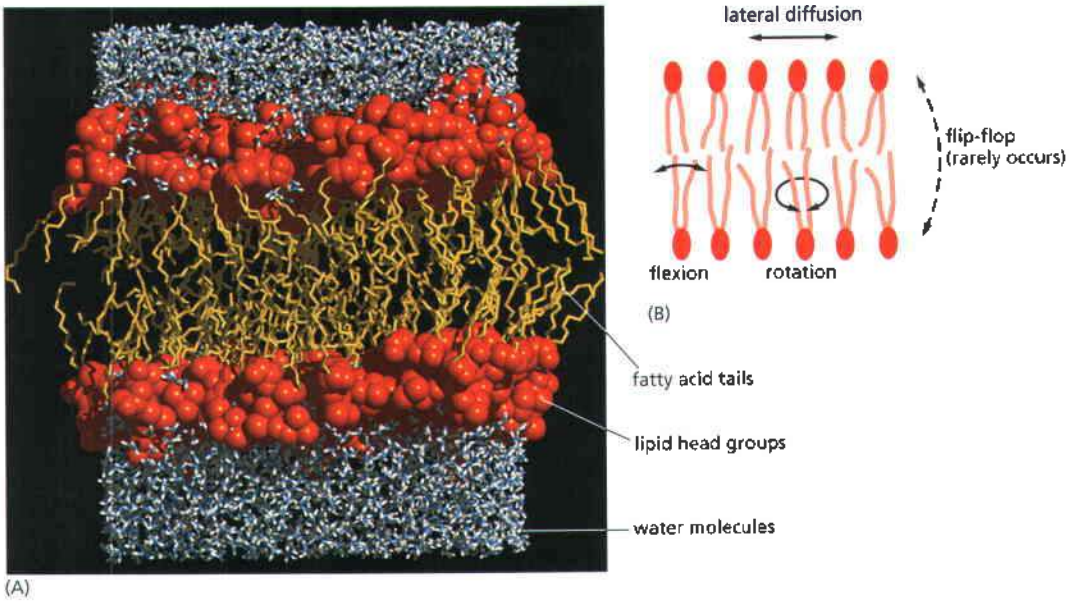


Figure 10–11 The mobility of phospholipid molecules in an artificial lipid bilayer. <CACA> Starting with a model of 100 phosphatidylcholine molecules arranged in a regular bilayer, a computer calculated the position of every atom after 300 picoseconds of simulated time. From these theoretical calculations (taking weeks of processor time in 1995), a model of the lipid bilayer emerges that accounts for almost all of the measurable properties of a synthetic lipid bilayer, such as its thickness, number of lipid molecules per membrane area, depth of water penetration, and unevenness of the two surfaces. Note that the tails in one monolayer can interact with those in the other monolayer, if the tails are long enough. (B) The different motions of a lipid molecule in a bilayer. (A, based on S.W. Chiu et al., *Biophys. J.* 69:1230–1245, 1995. With permission from the Biophysical Society.)

two-dimensional rigid crystalline (or gel) state at a characteristic freezing point. This change of state is called a *phase transition*, and the temperature at which it occurs is lower (that is, the membrane becomes more difficult to freeze) if the hydrocarbon chains are short or have double bonds. A shorter chain length reduces the tendency of the hydrocarbon tails to interact with one another, in both the same and opposite monolayer, and *cis*-double bonds produce kinks in the hydrocarbon chains that make them more difficult to pack together, so that the membrane remains fluid at lower temperatures (**Figure 10–12**). Bacteria, yeasts, and other organisms whose temperature fluctuates with that of their environment adjust the fatty acid composition of their membrane lipids to maintain a relatively constant fluidity. As the temperature falls, for instance, the cells of those organisms synthesize fatty acids with more *cis*-double bonds, and they avoid the decrease in bilayer fluidity that would otherwise result from the temperature drop.

Cholesterol modulates the properties of lipid bilayers. When mixed with phospholipids, it enhances the permeability-barrier properties of the lipid bilayer. It inserts into the bilayer with its hydroxyl group close to the polar head groups of the phospholipids, so that its rigid, platelike steroid rings interact with—and partly immobilize—those regions of the hydrocarbon chains closest to the polar head groups (see Figure 10–5). By decreasing the mobility of the first

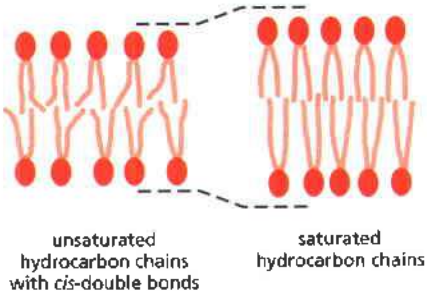


Figure 10–12 The influence of *cis*-double bonds in hydrocarbon chains. The double bonds make it more difficult to pack the chains together, thereby making the lipid bilayer more difficult to freeze. In addition, because the hydrocarbon chains of unsaturated lipids are more spread apart, lipid bilayers containing them are thinner than bilayers formed exclusively from saturated lipids.

Table 10–1 Approximate Lipid Compositions of Different Cell Membranes

LIPID	PERCENTAGE OF TOTAL LIPID BY WEIGHT					
	LIVER CELL PLASMA MEMBRANE	RED BLOOD CELL PLASMA MEMBRANE	MYELIN	MITOCHONDRION (INNER AND OUTER MEMBRANES)	ENDOPLASMIC RETICULUM	*E. COLI* BACTERIUM
Cholesterol	17	23	22	3	6	0
Phosphatidylethanolamine	7	18	15	28	17	70
Phosphatidylserine	4	7	9	2	5	trace
Phosphatidylcholine	24	17	10	44	40	0
Sphingomyelin	19	18	8	0	5	0
Glycolipids	7	3	28	trace	trace	0
Others	22	14	8	23	27	30

few CH_2 groups of the hydrocarbon chains of the phospholipid molecules, cholesterol makes the lipid bilayer less deformable in this region and thereby decreases the permeability of the bilayer to small water-soluble molecules. Although cholesterol tightens the packing of the lipids in a bilayer, it does not make membranes any less fluid. At the high concentrations found in most eucaryotic plasma membranes, cholesterol also prevents the hydrocarbon chains from coming together and crystallizing.

Table 10–1 compares the lipid compositions of several biological membranes. Note that bacterial plasma membranes are often composed of one main type of phospholipid and contain no cholesterol; their mechanical stability is enhanced by an overlying cell wall (see Figure 11–18). In archaea, lipids usually contain 20–25-carbon-long prenyl chains instead of fatty acids, prenyl and fatty acid chains are similarly hydrophobic and flexible (see Figure 10–20F). Thus, lipid bilayers can be built from molecules with similar features but different molecular designs. The plasma membranes of most eucaryotic cells are more varied than those of procaryotes and archaea, not only in containing large amounts of cholesterol but also in containing a mixture of different phospholipids.

Analysis of membrane lipids by mass spectrometry has revealed that the lipid composition of a typical cell membrane is much more complex than originally thought. According to these studies, membranes are composed of a bewildering variety of 500–1000 different lipid species. While some of this complexity reflects the combinatorial variation in head groups, hydrocarbon chain lengths, and desaturation of the major phospholipid classes, membranes also contain many structurally distinct minor lipids, at least some of which have important functions. The *inositol phospholipids*, for example, are present in small quantities but have crucial functions in guiding membrane traffic and in cell signaling (discussed in Chapters 13 and 15, respectively). Their local synthesis and destruction are regulated by a large number of enzymes, which create both small intracellular signaling molecules and lipid docking sites on membranes that recruit specific proteins from the cytosol, as we discuss later.

Despite Their Fluidity, Lipid Bilayers Can Form Domains of Different Compositions

Because a lipid bilayer is a two-dimensional fluid, we might expect most types of lipid molecules in it to be randomly distributed in their own monolayer. The van der Waals attractive forces between neighboring hydrocarbon tails are not selective enough to hold groups of phospholipid molecules together. With certain lipid mixtures, however, different lipids can come together transiently, creating a dynamic patchwork of different domains. In synthetic lipid bilayers composed of phosphatidylcholine, sphingomyelin, and cholesterol, van der Waals forces between the long and saturated hydrocarbon chains of the sphingomyelin molecules can be just strong enough to hold the adjacent molecules together transiently (Figure 10–13).

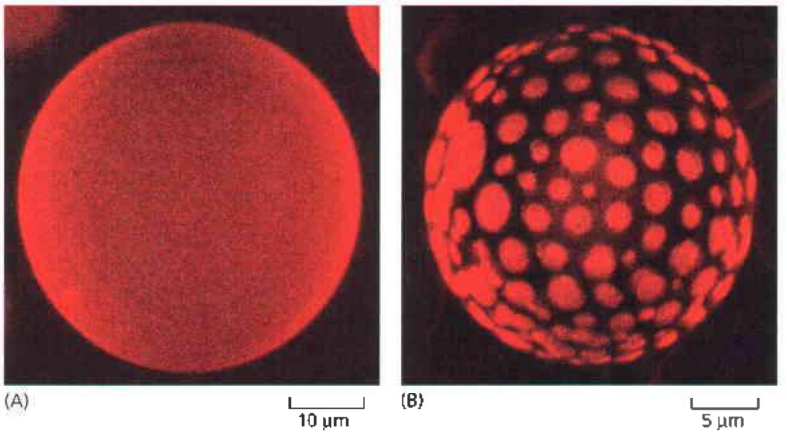


Figure 10–13 Lateral phase separation in artifical lipid bilayers. (A) Giant liposomes produced from a 1:1 mixture of phosphatidylcholine and spingomyelin form uniform bilayers, whereas (B) liposomes produced from a 1:1:1 mixture of phosphatidylcholine, spingomyelin, and cholesterol form bilayers with two immiscible phases. The liposomes are stained with trace concentrations of a fluorescent dye that preferentially partitions into one of the phases. The average size of the domains formed in these giant artificial liposomes is much larger than that expected in biological membranes, where rafts may be as small as a few nanometers in diameter. (A, from N. Kahya et al., *J. Struct. Biol.* 147:77–89, 2004. With permission from Elsevier; B, courtesy of Petra Schwille.)

There has been a long debate among scientists whether the lipid molecules in the plasma membrane of animal cells can transiently assemble into specialized domains, called **lipid rafts**. Certain specialized regions of the plasma membrane, such as the *caveolae* involved in endocytosis (discussed in Chapter 13), are enriched in sphingolipids and cholesterol, and it is thought that the specific proteins that assemble there help stabilize these rafts. Because the hydrocarbon chains of sphingolipids are longer and straighter than those of other membrane lipids, raft domains are thicker than other parts of the bilayer (see Figure 10–12) and better accommodate certain membrane proteins (**Figure 10–14**). Thus, the lateral segregation of proteins and of lipids into raft domains would, in principle, be a mutually stabilizing process. In this way, lipid rafts could help organize membrane proteins—concentrating them either for transport in membrane vesicles (discussed in Chapter 13) or for working together in protein assemblies, as when they convert extracellular signals into intracellular ones (discussed in Chapter 15).

Lipid Droplets Are Surrounded by a Phospholipid Monolayer

Most cells store an excess of lipids in **lipid droplets**, from where they can be retrieved as building blocks for membrane synthesis or as a food source. Fat cells, also called adipocytes, are specialized for lipid storage (see Figure 14–34). They contain vast numbers of large lipid droplets, from which fatty acids can be liberated on demand and exported to other cells through the bloodstream. Lipid droplets store neutral lipids, such as triacylglycerides and cholesterol esters, which are synthesized from fatty acids and cholesterol by enzymes in the

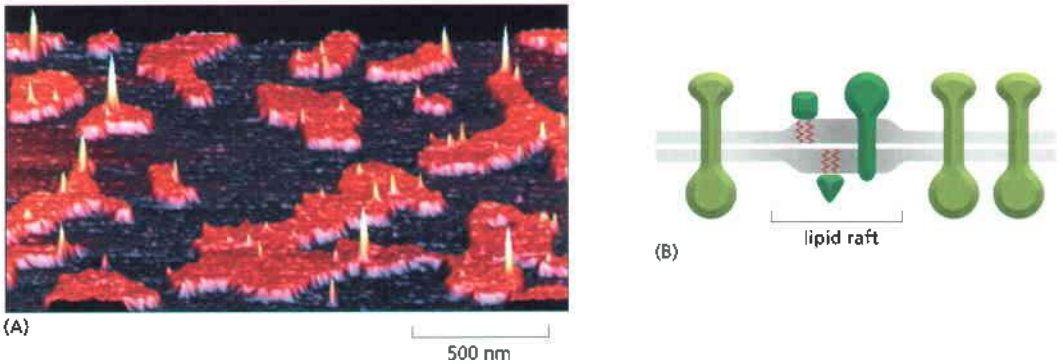


Figure 10–14 The effects of lipid rafts in artificial lipid bilayers. (A) The surface contours of a synthetic bilayer containing lipid rafts, analyzed by atomic force microscopy. Note that the raft areas, shown in *orange*, are thicker than the rest of the bilayer; as in Figure 10–13, the rafts primarily contain sphingomyelin and cholesterol. The sharp, yellow spikes are incorporated protein molecules, which are attached to the bilayer by a glycosylphosphatidyl-inositol (GPI) anchor (illustrated in Figure 10–19, example 6), and preferentially partition into the raft domains. (B) Because of both their increased thickness and lipid composition, rafts are thought to concentrate specific membrane proteins *(dark green)*. (A, from D.E. Saslowsky et al., *J. Biol. Chem.* 277:26966–26970, 2002. With permission from the American Society for Biochemistry and Molecular Biology.)

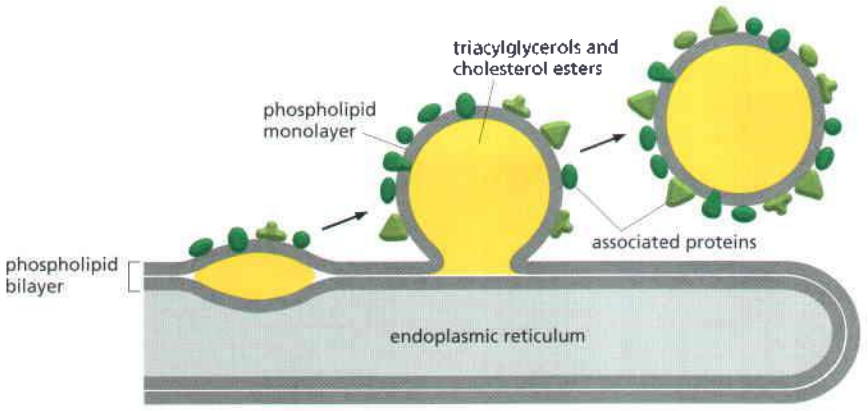


Figure 10–15 A model for the formation of lipid droplets. Neutral lipids are deposited between the two monolayers of the endoplasmic reticulum membrane. There, they aggregate into a three-dimensional droplet, which buds and pinches off from the endoplasmic reticulum membrane as a unique organelle, surrounded by a single monolayer of phospholipids and associated proteins. (Adapted from S. Martin and R.G. Parton, *Nat. Rev. Mol. Cell Biol.* 7:373–378, 2006. With permission from Macmillan Publishers Ltd.)

endoplasmic reticulum membrane. Because these lipids do not contain hydrophilic head groups, they are exclusively hydrophobic molecules, which aggregate into three-dimensional droplets rather than into bilayers.

Lipid droplets are unique organelles because they are surrounded by a single monolayer of phospholipids, which contains a large variety of proteins. Some of the proteins are enzymes involved in lipid metabolism, but the functions of most are unknown. Lipid droplets form rapidly when cells are exposed to high concentrations of fatty acids. They form from discrete regions of the endoplasmic reticulum membrane where many enzymes of lipid metabolism are concentrated. **Figure 10–15** shows one model of how lipid droplets may form and acquire their surrounding monolayer of phospholipids and proteins.

The Asymmetry of the Lipid Bilayer Is Functionally Important

The lipid compositions of the two monolayers of the lipid bilayer in many membranes are strikingly different. In the human red blood cell membrane, for example, almost all of the phospholipid molecules that have choline—$(CH_3)_3N^+CH_2CH_2OH$—in their head group (phosphatidylcholine and sphingomyelin) are in the outer monolayer, whereas almost all that contain a terminal primary amino group (phosphatidylethanolamine and phosphatidylserine) are in the inner monolayer (**Figure 10–16**). Because the negatively charged phosphatidylserine is located in the inner monolayer, there is a significant difference in charge between the two halves of the bilayer. We discuss in Chapter 12 how membrane-bound phospholipid translocators generate and maintain lipid asymmetry.

Lipid asymmetry is functionally important, especially in converting extracellular signals into intracellular ones (discussed in Chapter 15). Many cytosolic proteins bind to specific lipid head groups found in the cytosolic monolayer of the lipid bilayer. The enzyme *protein kinase C (PKC)*, for example, is activated in response to various extracellular signals. It binds to the cytosolic face of the plasma membrane, where phosphatidylserine is concentrated, and requires this negatively charged phospholipid for its activity.

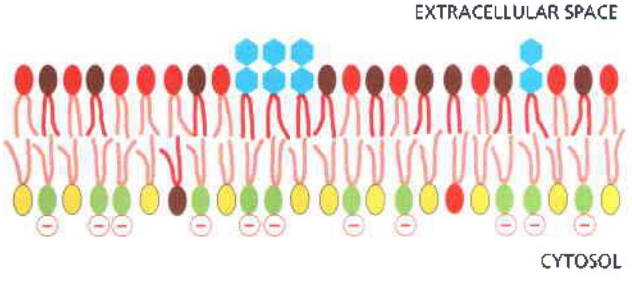


Figure 10–16 The asymmetrical distribution of phospholipids and glycolipids in the lipid bilayer of human red blood cells. The colors used for the phospholipid head groups are those introduced in Figure 10–3. In addition, glycolipids are drawn with hexagonal polar head groups *(blue)*. Cholesterol (not shown) is thought to be distributed roughly equally in both monolayers.

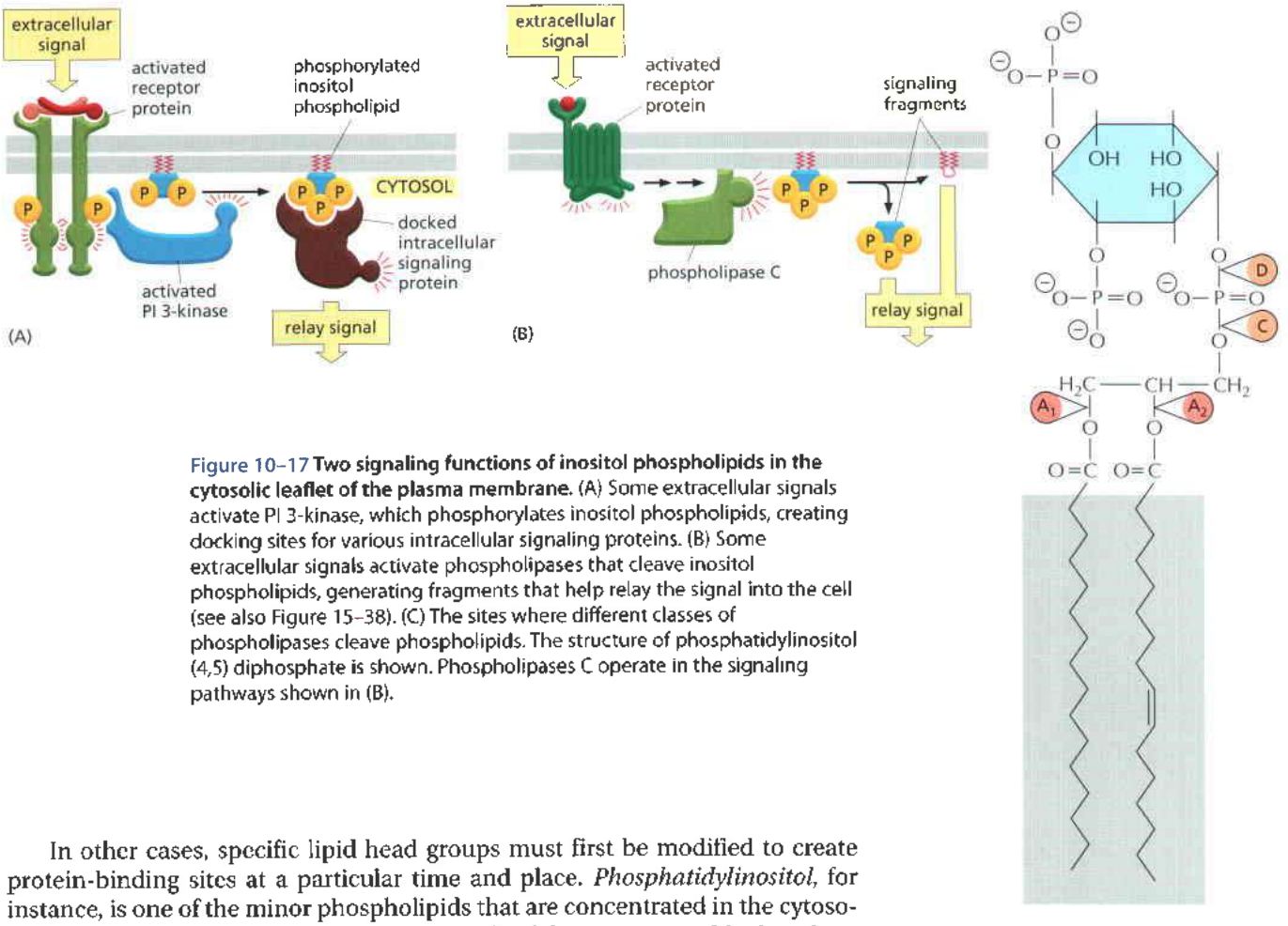


Figure 10–17 **Two signaling functions of inositol phospholipids in the cytosolic leaflet of the plasma membrane.** (A) Some extracellular signals activate PI 3-kinase, which phosphorylates inositol phospholipids, creating docking sites for various intracellular signaling proteins. (B) Some extracellular signals activate phospholipases that cleave inositol phospholipids, generating fragments that help relay the signal into the cell (see also Figure 15–38). (C) The sites where different classes of phospholipases cleave phospholipids. The structure of phosphatidylinositol (4,5) diphosphate is shown. Phospholipases C operate in the signaling pathways shown in (B).

In other cases, specific lipid head groups must first be modified to create protein-binding sites at a particular time and place. *Phosphatidylinositol,* for instance, is one of the minor phospholipids that are concentrated in the cytosolic monolayer of cell membranes. Various lipid kinases can add phosphate groups at distinct positions in the inositol ring, creating binding sites that recruit specific proteins from the cytosol to the membrane. An important example of such a lipid kinase is *phosphoinositide 3-kinase (PI 3-kinase),* which is activated in response to extracellular signals and helps to recruit specific intracellular signaling proteins to the cytosolic face of the plasma membrane (**Figure 10–17**A). Similar lipid kinases phosphorylate inositol phospholipids in intracellular membranes and thereby help to recruit proteins that guide membrane transport.

Phospholipids in the plasma membrane are used in yet another way to convert extracellular signals into intracellular ones. The plasma membrane contains various *phospholipases* that are activated by extracellular signals to cleave specific phospholipid molecules, generating fragments of these molecules that act as short-lived intracellular mediators. *Phospholipase C,* for example, cleaves an inositol phospholipid in the cytosolic monolayer of the plasma membrane to generate two fragments, one of which remains in the membrane and helps activate protein kinase C, while the other is released into the cytosol and stimulates the release of Ca^{2+} from the endoplasmic reticulum (Figure 10–17B–C).

Animals exploit the phospholipid asymmetry of their plasma membranes to distinguish between live and dead cells. When animal cells undergo apoptosis (a form of programmed cell death, discussed in Chapter 18), phosphatidylserine, which is normally confined to the cytosolic monolayer of the plasma membrane lipid bilayer, rapidly translocates to the extracellular monolayer. The phosphatidylserine exposed on the cell surface signals neighboring cells, such as macrophages, to phagocytose the dead cell and digest it. The translocation of the phosphatidylserine in apoptotic cells is thought to occur by two mechanisms:

1. The phospholipid translocator that normally transports this lipid from the noncytosolic monolayer to the cytosolic monolayer is inactivated.
2. A "scramblase" that transfers phospholipids nonspecifically in both directions between the two monolayers is activated.

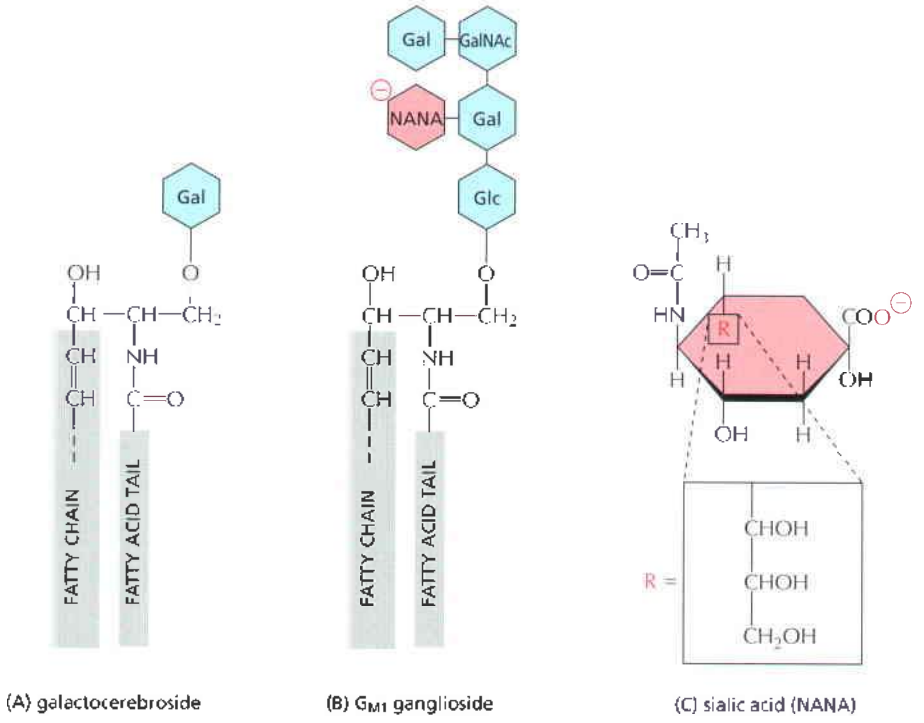


Figure 10–18 Glycolipid molecules. (A) Galactocerebroside is called a *neutral glycolipid* because the sugar that forms its head group is uncharged. (B) A ganglioside always contains one or more negatively charged sialic acid residues (also called *N*-acetylneuraminic acid, or NANA), whose structure is shown in (C). Whereas in bacteria and plants almost all glycolipids are derived from glycerol, as are most phospholipids, in animal cells almost all glycolipids are based on sphingosine, as is the case for sphingomyelin (see Figure 10–3). Gal = galactose; Glc = glucose, GalNAc = *N*-acetylgalactosamine; these three sugars are uncharged.

Glycolipids Are Found on the Surface of All Plasma Membranes

Sugar-containing lipid molecules called **glycolipids**, found exclusively in the noncytosolic monolayer of the lipid bilayer, have the most extreme asymmetry in their membrane distribution. In animal cells they are made from sphingosine, just like sphingomyelin. These intriguing molecules tend to self-associate, partly through hydrogen bonds between their sugars and partly through van der Waals forces between their long and straight hydrocarbon chains, and they may preferentially partition into lipid rafts. The asymmetric distribution of glycolipids in the bilayer results from the addition of sugar groups to the lipid molecules in the lumen of the Golgi apparatus. Thus, the compartment in which they are manufactured is topologically equivalent to the exterior of the cell (discussed in Chapter 12). As they are delivered to the plasma membrane, the sugar groups are exposed at the cell surface (see Figure 10–16), where they have important roles in interactions of the cell with its surroundings.

Glycolipids probably occur in all animal cell plasma membranes, where they generally constitute about 5% of the lipid molecules in the outer monolayer. They are also found in some intracellular membranes. The most complex of the glycolipids, the **gangliosides**, contain oligosaccharides with one or more sialic acid residues, which give gangliosides a net negative charge (**Figure 10–18**). The most abundant of the more than 40 different gangliosides that have been identified are in the plasma membrane of nerve cells, where gangliosides constitute 5–10% of the total lipid mass; they are also found in much smaller quantities in other cell types.

Hints as to the functions of glycolipids come from their localization. In the plasma membrane of epithelial cells, for example, glycolipids are confined to the exposed apical surface, where they may help to protect the membrane against the harsh conditions frequently found there (such as low pH and high concentrations of degradative enzymes). Charged glycolipids, such as gangliosides, may be important because of their electrical effects: their presence alters the electrical field across the membrane and the concentrations of ions—especially Ca^{2+}—at the membrane surface. Glycolipids are also thought to function in cell-recognition processes, in which membrane-bound carbohydrate-binding proteins *(lectins)* bind to the sugar groups on both glycolipids and glycoproteins in the process of cell–cell adhesion (discussed in Chapter 19). Surprisingly, however, mutant mice that are deficient in all of their complex gangliosides

show no obvious abnormalities, although the males cannot transport testosterone normally in the testes and are consequently sterile.

Whatever their normal function, some glycolipids provide entry points for certain bacterial toxins. The ganglioside G_{M1} (see Figure 10–18), for example, acts as a cell-surface receptor for the bacterial toxin that causes the debilitating diarrhea of cholera. Cholera toxin binds to and enters only those cells that have G_{M1} on their surface, including intestinal epithelial cells. Its entry into a cell leads to a prolonged increase in the concentration of intracellular cyclic AMP (discussed in Chapter 15), which in turn causes a large efflux of Na^+ and water into the intestine.

Summary

Biological membranes consist of a continuous double layer of lipid molecules in which membrane proteins are embedded. This lipid bilayer is fluid, with individual lipid molecules able to diffuse rapidly within their own monolayer. The membrane lipid molecules are amphiphilic. When placed in water they assemble spontaneously into bilayers, which form sealed compartments.

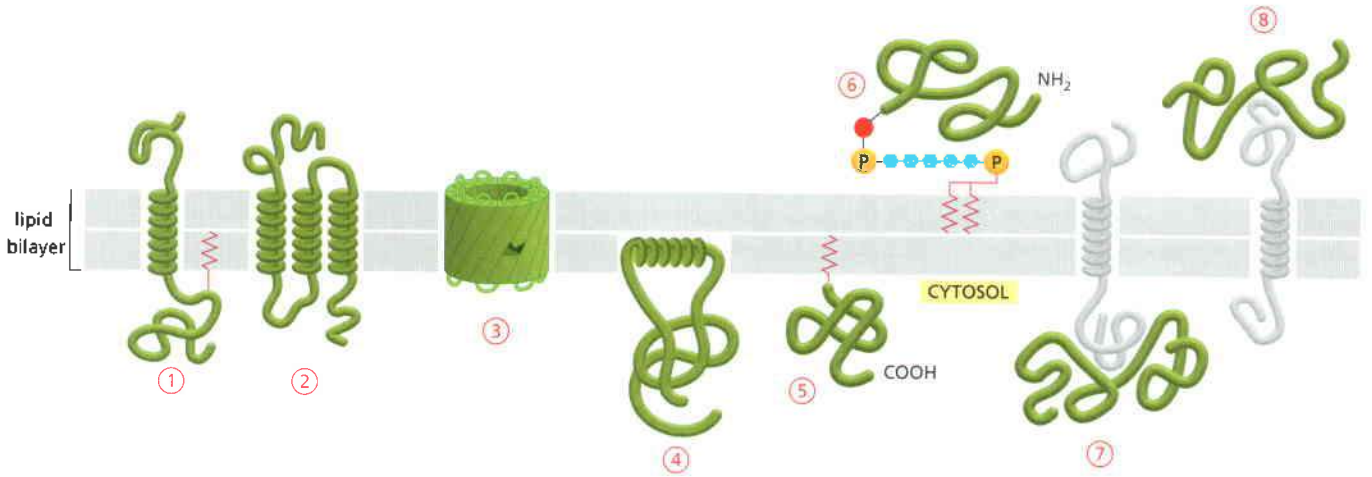
Cells contain 500–1000 different lipid species. There are three major classes of membrane lipids—phospholipids, cholesterol, and glycolipids—and hundreds of minor classes. The lipid compositions of the inner and outer monolayers are different, reflecting the different functions of the two faces of a cell membrane. Different mixtures of lipids are found in the membranes of cells of different types, as well as in the various membranes of a single eucaryotic cell. Inositol phospholids are a minor class of phospholipids, which in the cytosolic leaflet of the plasma membrane lipid bilayer play an important part in cell signaling: in response to extracellular signals, specific lipid kinases phosphorylate the head groups of these lipids to form docking sites for cytosolic signaling proteins, whereas specific phospholipases cleave certain inositol phospholipids to generate small intracellular signaling molecules.

MEMBRANE PROTEINS

Although the lipid bilayer provides the basic structure of biological membranes, the membrane proteins perform most of the membrane's specific tasks and therefore give each type of cell membrane its characteristic functional properties. Accordingly, the amounts and types of proteins in a membrane are highly variable. In the myelin membrane, which serves mainly as electrical insulation for nerve cell axons, less than 25% of the membrane mass is protein. By contrast, in the membranes involved in ATP production (such as the internal membranes of mitochondria and chloroplasts), approximately 75% is protein. A typical plasma membrane is somewhere in between, with protein accounting for about half of its mass. Because lipid molecules are small compared with protein molecules, there are always many more lipid molecules than protein molecules in cell membranes—about 50 lipid molecules for each protein molecule in cell membranes that are 50% protein by mass. Membrane proteins vary widely in structure and in the way they associate with the lipid bilayer, which reflects their diverse functions.

Membrane Proteins Can Be Associated with the Lipid Bilayer in Various Ways

Figure 10–19 shows the different ways in which membrane proteins can associate with the membrane. Many extend through the lipid bilayer, with part of their mass on either side (Figure 10–19, examples 1, 2, and 3). Like their lipid neighbors, these **transmembrane proteins** are amphiphilic, having hydrophobic and hydrophilic regions. Their hydrophobic regions pass through the membrane and interact with the hydrophobic tails of the lipid molecules in the interior of the bilayer, where they are sequestered away from water. Their hydrophilic regions are exposed to water on either side of the membrane. The covalent attachment of a fatty acid chain that inserts into the cytosolic monolayer of the



lipid bilayer increases the hydrophobicity of some of these transmembrane proteins (see Figure 10–19, example 1).

Other membrane proteins are located entirely in the cytosol and are associated with the cytosolic monolayer of the lipid bilayer, either by an amphiphilic α helix exposed on the surface of the protein (Figure 10–19, example 4) or by one or more covalently attached lipid chains (Figure 10–19, example 5). Yet other membrane proteins are entirely exposed at the external cell surface, being attached to the lipid bilayer only by a covalent linkage (via a specific oligosaccharide) to phosphatidylinositol in the outer lipid monolayer of the plasma membrane (Figure 10–19, example 6).

The lipid-linked proteins in example 5 in Figure 10–19 are made as soluble proteins in the cytosol and are subsequently anchored to the membrane by the covalent attachment of a lipid group. The proteins in example 6, however, are made as single-pass transmembrane proteins in the endoplasmic reticulum (ER). While still in the ER, the transmembrane segment of the protein is cleaved off and a **glycosylphosphatidylinositol (GPI) anchor** is added, leaving the protein bound to the noncytosolic surface of the membrane solely by this anchor (discussed in Chapter 12). Transport vesicles eventually deliver the protein to the plasma membrane (discussed in Chapter 13). Proteins bound to the plasma membrane by a GPI anchor can be readily distinguished by the use of an enzyme called phosphatidylinositol-specific phospholipase C. This enzyme cuts these proteins free from their anchors, thereby releasing them from the membrane.

Some membrane proteins do not extend into the hydrophobic interior of the lipid bilayer at all; they are instead bound to either face of the membrane by noncovalent interactions with other membrane proteins (Figure 10–19, examples 7 and 8). Many of the proteins of this type can be released from the membrane by relatively gentle extraction procedures, such as exposure to solutions of very high or low ionic strength or of extreme pH, which interfere with protein–protein interactions but leave the lipid bilayer intact; these proteins are referred to as **peripheral membrane proteins.** Transmembrane proteins and many proteins held in the bilayer by lipid groups or hydrophobic polypeptide regions that insert into the hydrophobic core of the lipid bilayer cannot be released in these ways. These proteins are called **integral membrane proteins**.

Figure 10–19 Various ways in which membrane proteins associate with the lipid bilayer. Most transmembrane proteins are thought to extend across the bilayer as (1) a single α helix, (2) as multiple α helices, or (3) as a rolled-up β sheet (a β barrel). Some of these "single-pass" and "multipass" proteins have a covalently attached fatty acid chain inserted in the cytosolic lipid monolayer (1). Other membrane proteins are exposed at only one side of the membrane. (4) Some of these are anchored to the cytosolic surface by an amphiphilic α helix that partitions into the cytosolic monolayer of the lipid bilayer through the hydrophobic face of the helix. (5) Others are attached to the bilayer solely by a covalently attached lipid chain—either a fatty acid chain or a prenyl group (see Figure 10–20)—in the cytosolic monolayer or, (6) via an oligosaccharide linker, to phosphatidylinositol in the noncytosolic monolayer—called a GPI anchor. (7, 8) Finally, many proteins are attached to the membrane only by noncovalent interactions with other membrane proteins. The way in which the structure in (5) is formed is illustrated in Figure 10–20, while the way in which the GPI anchor shown in (6) is formed is illustrated in Figure 12–56. The details of how membrane proteins become associated with the lipid bilayer are discussed in Chapter 12.

Lipid Anchors Control the Membrane Localization of Some Signaling Proteins

How a membrane protein is associated with the lipid bilayer reflects the function of the protein. Only transmembrane proteins can function on both sides of the bilayer or transport molecules across it. Cell-surface receptors, for example, are transmembrane proteins that bind signal molecules in the extracellular space and generate different intracellular signals on the opposite side of the plasma membrane. To transfer small hydrophilic molecules across a membrane,

a membrane transport protein must provide a path for the molecules to cross the hydrophobic permeability barrier of the lipid bilayer; the molecular architecture of multipass membrane proteins is ideally suited for this task, as we discuss in Chapter 11.

Proteins that function on only one side of the lipid bilayer, by contrast, are often associated exclusively with either the lipid monolayer or a protein domain on that side. Some intracellular signaling proteins, for example, that are involved in converting extracellular signals into intracellular ones are bound to the cytosolic half of the plasma membrane by one or more covalently attached lipid groups, which can be fatty acid chains or *prenyl groups* (**Figure 10–20**). In some cases, myristic acid, a saturated 14-carbon fatty acid, is added to the N-terminal amino group of the protein during its synthesis on the ribosome. All members of the *Src family* of cytoplasmic protein tyrosine kinases (discussed in Chapter 15) are myristoylated in this way. Membrane attachment through a single lipid anchor is not very strong, however, and a second lipid group is often added to anchor proteins more firmly to a membrane. For most Src kinases, the second lipid modification is the attachment of palmitic acid, a saturated 16-carbon fatty acid, to a cysteine side chain of the protein. This modification occurs in response to an extracellular signal and helps recruit the kinases to the plasma membrane. When the signaling pathway is turned off, the palmitic acid is removed, allowing the kinase to return to the cytosol. Other intracellular signaling proteins, such as the Ras family small GTPases (discussed in Chapter 15), use a combination of prenyl group and palmitic acid attachment to recruit the proteins to the plasma membrane.

In Most Transmembrane Proteins the Polypeptide Chain Crosses the Lipid Bilayer in an α-Helical Conformation

A transmembrane protein always has a unique orientation in the membrane. This reflects both the asymmetric manner in which it is inserted into the lipid

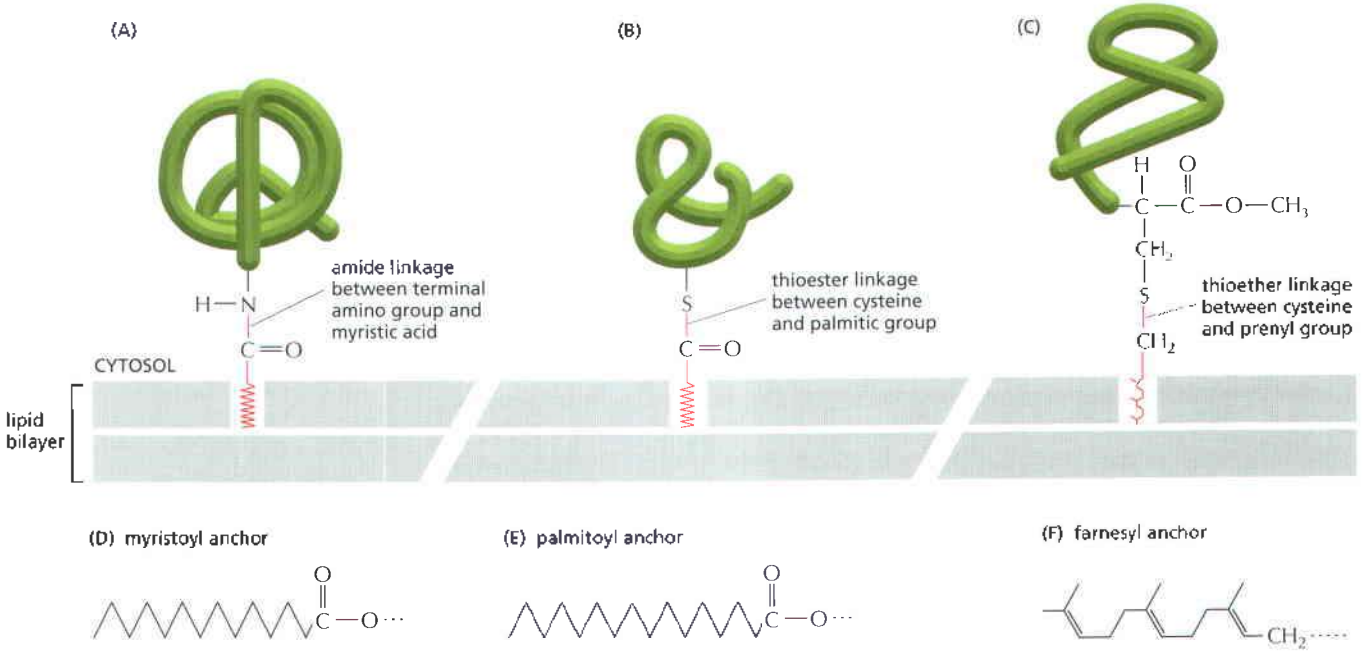


Figure 10–20 **Membrane protein attachment by a fatty acid chain or a prenyl group.** The covalent attachment of either type of lipid can help localize a water-soluble protein to a membrane after its synthesis in the cytosol. (A) A fatty acid chain (myristic acid) is attached via an amide linkage to an N-terminal glycine. (B) A fatty acid chain (palmitic acid) is attached via a thioester linkage to a cysteine. (C) A prenyl group (either farnesyl or a longer geranylgeranyl group) is attached via a thioether linkage to a cysteine residue that is initially located four residues from the protein's C-terminus. After prenylation, the terminal three amino acids are cleaved off, and the new C-terminus is methylated before insertion of the anchor into the membrane (not shown). The structures of the lipid anchors are shown below: (D) a myristoyl anchor (a 14-carbon saturated fatty acid chain), (E) a palmitoyl anchor (a 16-carbon saturated fatty acid chain), and (F) a farnesyl anchor (a 15-carbon unsaturated hydrocarbon chain).

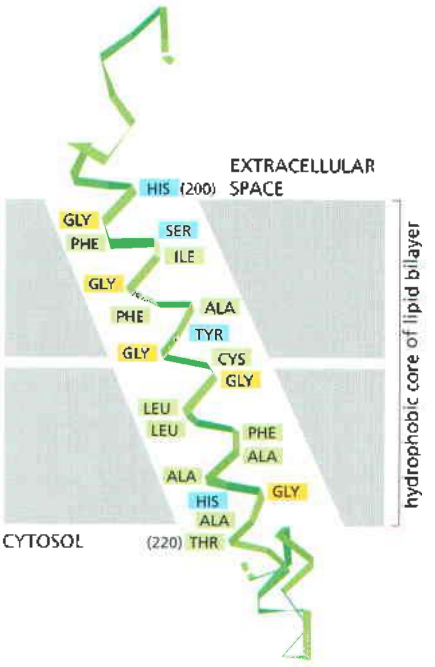


Figure 10–21 A segment of a transmembrane polypeptide chain crossing the lipid bilayer as an α helix. <GTAG> Only the α-carbon backbone of the polypeptide chain is shown, with the hydrophobic amino acids in *green* and *yellow*. The polypeptide segment shown is part of the bacterial photosynthetic reaction center illustrated in Figure 10–34, the structure of which was determined by x-ray diffraction. (Based on data from J. Deisenhofer et al., *Nature* 318:618–624, 1985, and H. Michel et al., *EMBO J.* 5:1149–1158, 1986. All with permission from Macmillan Publishers Ltd.)

bilayer in the ER during its biosynthesis (discussed in Chapter 12) and the different functions of its cytosolic and noncytosolic domains. These domains are separated by the membrane-spanning segments of the polypeptide chain, which contact the hydrophobic environment of the lipid bilayer and are composed largely of amino acids with nonpolar side chains. Because the peptide bonds themselves are polar and because water is absent, all peptide bonds in the bilayer are driven to form hydrogen bonds with one another. The hydrogen-bonding between peptide bonds is maximized if the polypeptide chain forms a regular α helix as it crosses the bilayer, and this is how most membrane-spanning segments of polypeptide chains traverse the bilayer (**Figure 10–21**).

In **single-pass transmembrane proteins**, the polypeptide chain crosses only once (see Figure 10–19, example 1), whereas in **multipass transmembrane proteins**, the polypeptide chain crosses multiple times (see Figure 10–19, example 2). An alternative way for the peptide bonds in the lipid bilayer to satisfy their hydrogen-bonding requirements is for multiple transmembrane strands of a polypeptide chain to be arranged as a β sheet that is rolled up into a closed barrel (a so-called *β barrel*; see Figure 10–19, example 3). This form of multipass transmembrane structure is seen in the *porin proteins* that we discuss later.

Rapid progress in the x-ray crystallography of membrane proteins has enabled us to determine the three-dimensional structure of many of them. The structures confirm that it is often possible to predict from the protein's amino acid sequence which parts of the polypeptide chain extend across the lipid bilayer. Segments containing about 20–30 amino acids with a high degree of hydrophobicity are long enough to span a lipid bilayer as an α helix, and they can often be identified in *hydropathy plots* (**Figure 10–22**). From such plots, it is estimated that about 20% of the kind of an organism's proteins are transmembrane proteins, emphasizing their importance. Hydropathy plots cannot identify the membrane-spanning segments of a β barrel, as 10 amino acids or fewer are sufficient to traverse a lipid bilayer as an extended β strand and only every other amino acid side chain is hydrophobic.

The strong drive to maximize hydrogen-bonding in the absence of water means that a polypeptide chain that enters the bilayer is likely to pass entirely through it before changing direction, since chain bending requires a loss of regular hydrogen-bonding interactions. But multipass membrane proteins can also contain regions that fold into the membrane from either side, squeezing into spaces between transmembrane α helices without contacting the hydrophobic core of the lipid bilayer. Because such regions of the polypeptide chain interact only with other polypeptide regions, they do not need to maximize hydrogen-bonding; they can therefore have a variety of secondary structures, including helices that extend only part way across the lipid bilayer (**Figure 10–23**). Such regions are important for the function of some membrane proteins, including the K^+ and water channels; the regions contribute to the walls of the pores traversing the membrane and confer substrate specificity on the channels, as we discuss in Chapter 11. These regions cannot be identified in hydropathy plots and are only revealed by x-ray crystallography, electron diffraction (a technique similar to x-ray diffraction but performed on two-dimensional arrays of proteins), or NMR studies of the protein's three-dimensional structure.

Transmembrane α Helices Often Interact with One Another

The transmembrane α helices of many single-pass membrane proteins do not contribute to the folding of the protein domains on either side of the membrane.

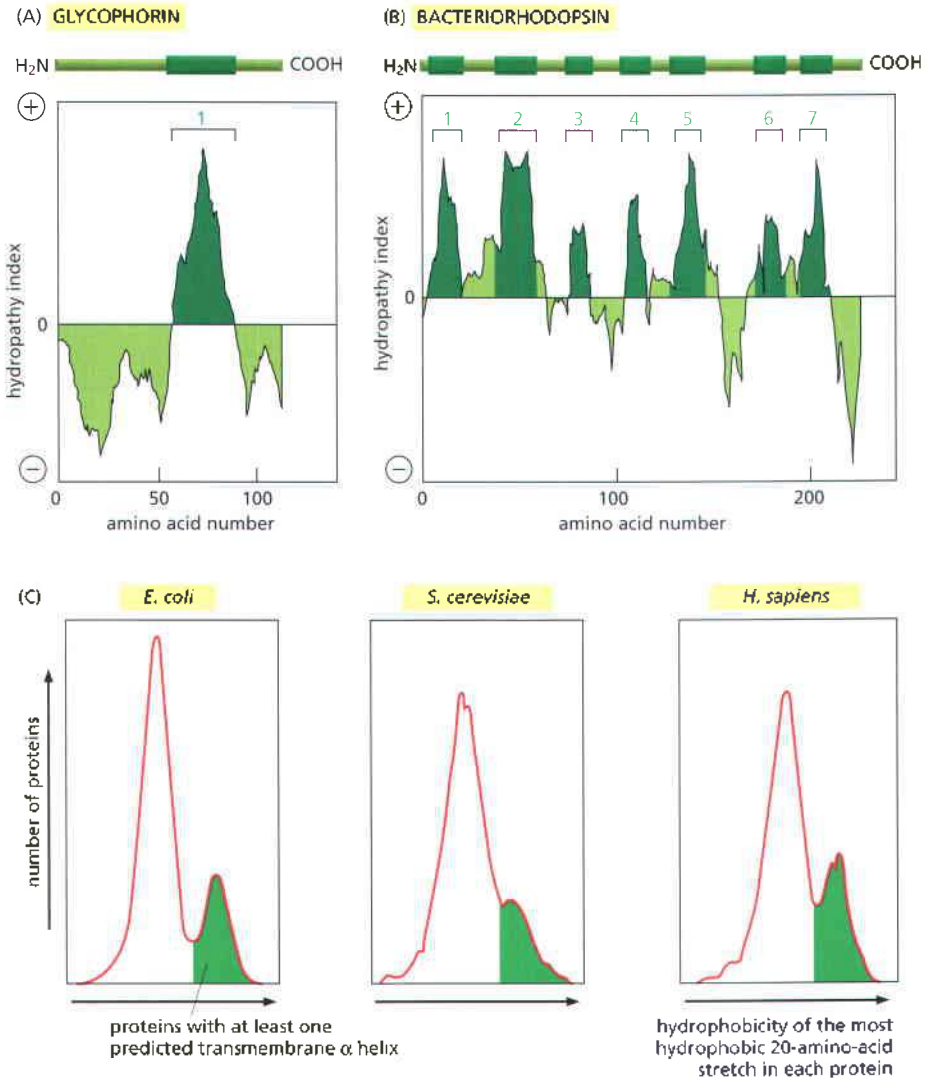


Figure 10–22 Using hydropathy plots to localize potential α-helical membrane-spanning segments in a polypeptide chain. The free energy needed to transfer successive segments of a polypeptide chain from a nonpolar solvent to water is calculated from the amino acid composition of each segment using data obtained from model compounds. This calculation is made for segments of a fixed size (usually around 10–20 amino acids), beginning with each successive amino acid in the chain. The "hydropathy index" of the segment is plotted on the Y axis as a function of its location in the chain. A positive value indicates that free energy is required for transfer to water (i.e., the segment is hydrophobic), and the value assigned is an index of the amount of energy needed. Peaks in the hydropathy index appear at the positions of hydrophobic segments in the amino acid sequence. (A and B) Two examples of membrane proteins discussed later in this chapter are shown. Glycophorin (A) has a single membrane-spanning α helix and one corresponding peak in the hydropathy plot. Bacteriorhodopsin (B) has seven membrane-spanning α helices and seven corresponding peaks in the hydropathy plot. (C) The proportion of predicted membrane proteins encoded by the genomes of *E. coli*, *S. cerevisiae*, and human. The area shaded in *green* indicates the fraction of proteins that contain at least one predicted transmembrane helix. The data for *E. coli* and *S. cerevisiae* represent the whole genome; the data for human represent only part of the genome; in each case, the area under the curve is proportional to the number of genes analyzed. (A, adapted from D. Eisenberg, *Annu. Rev. Biochem.* 53:595–624, 1984. With permission from Annual Reviews; C, adapted from D. Boyd et al., *Protein Sci.* 7:201–205, 1998. With permission from The Protein Society.)

As a consequence, it is often possible to engineer cells to produce the cytosolic or extracellular domains of these proteins as water-soluble protein. This approach has been invaluable to study the structure and function of these domains, especially of those in transmembrane receptor proteins (discussed in Chapter 15). A transmembrane α helix, even in a single-pass membrane protein, however, often does more than just anchor the protein to the lipid bilayer. Many single-pass membrane proteins form homodimers, which are held together by strong and highly specific interactions between the two transmembrane α helices; the sequence of the hydrophobic amino acids of these helices contains the information that directs the protein–protein interaction.

Similarly, the transmembrane α helices in multipass membrane proteins occupy specific positions in the folded protein structure that are determined by interactions between the neighboring helices. These interactions are crucial for the structure and function of the many channels and transporters that move molecules across lipid bilayers. In many cases, one can use proteases to cut the loops of the polypeptide chain that link the transmembrane α helices on either side of the bilayer and the helices stay together and function normally. In some

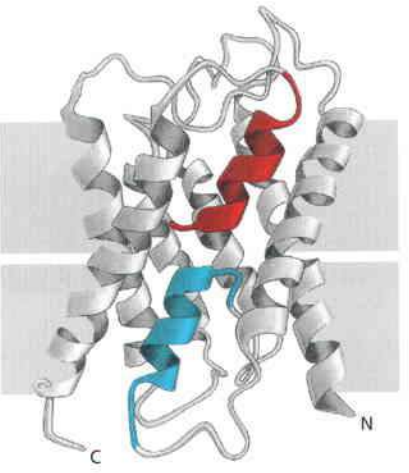


Figure 10–23 Two α helices in the aquaporin water channel, each of which spans only halfway through the lipid bilayer. In the membrane, the protein forms a tetramer of four such two-helix segments, such that the colored surface shown here is buried at an interface formed by protein–protein interactions. The mechanism by which the channel allows the passage of water molecules across the lipid bilayer is discussed in more detail in Chapter 11.

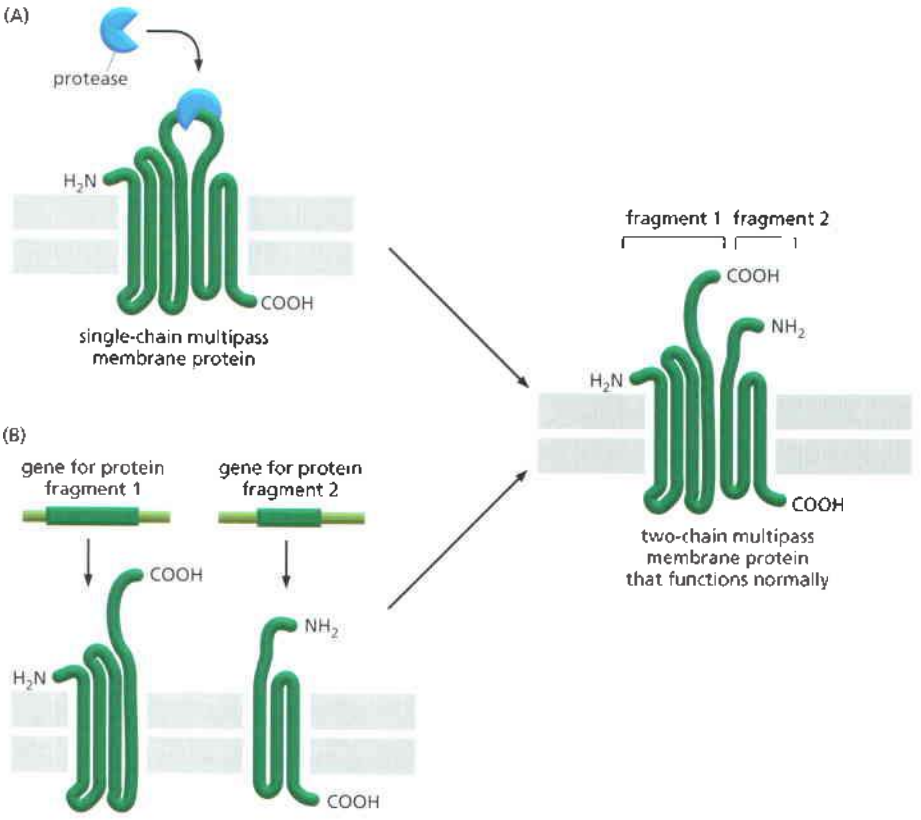


Figure 10–24 Converting a single-chain multipass protein into a two-chain multipass protein. (A) Proteolytic cleavage of one loop to create two fragments that stay together and function normally. (B) Expression of the same two fragments from separate genes gives rise to a similar protein that functions normally.

cases, one can even express engineered genes encoding separate pieces of a multipass protein in living cells, and one finds that the separate pieces assemble properly to form a functional transmembrane protein (**Figure 10–24**), emphasizing the exquisite specificity with which transmembrane α helices can interact.

In multipass membrane proteins, neighboring transmembrane helices in the folded structure of the protein shield many of the transmembrane helices from the membrane lipids. Why, then, are these shielded helices nevertheless composed primarily of hydrophobic amino acids? The answer lies in the way in which multipass proteins are integrated into the membrane during their biosynthesis. As we discuss in Chapter 12, transmembrane α helices are inserted into the lipid bilayer sequentially by a protein translocator. After leaving the translocator, each helix is transiently surrounded by lipids in the bilayer, which requires that the helix be hydrophobic. It is only as the protein folds up into its final structure that contacts are made between adjacent helices and protein–protein contacts replace some of the protein–lipid contacts (**Figure 10–25**).

Some β Barrels Form Large Transmembrane Channels

Multipass transmembrane proteins that have their transmembrane segments arranged as a *β barrel* rather than as an α helix are comparatively rigid and tend to crystallize readily. Thus, some of them were among the first multipass

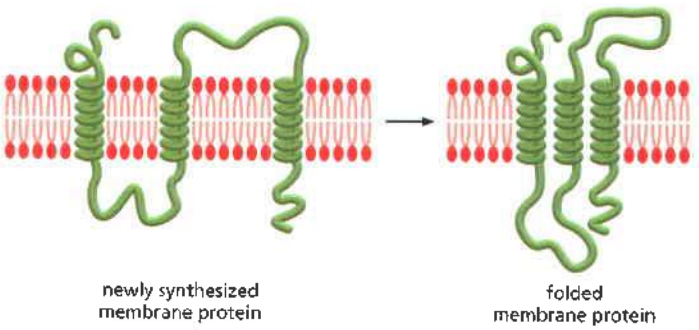


Figure 10–25 Steps in the folding of a multipass transmembrane protein. When the newly synthesized transmembrane α helices are released into the lipid bilayer, they are initially surrounded by lipid molecules. As the protein folds, contacts between the helices displace some of the lipid molecules surrounding the helices.

membrane protein structures to be determined by x-ray crystallography. The number of β strands in a β barrel varies widely, from as few as 8 strands to as many as 22 (**Figure 10–26**).

β barrel proteins are abundant in the outer membrane of mitochondria, chloroplasts, and many bacteria. Some are pore-forming proteins, which create water-filled channels that allow selected small hydrophilic molecules to cross the lipid bilayer of the bacterial outer membrane. The porins are well-studied examples (example 3 in Figure 10–26). The porin barrel is formed from a 16-strand, antiparallel β sheet, which is sufficiently large to roll up into a cylindrical structure. Polar amino acid side chains line the aqueous channel on the inside, while nonpolar side chains project from the outside of the barrel to interact with the hydrophobic core of the lipid bilayer. Loops of the polypeptide chain often protrude into the lumen of the channel, narrowing it so that only certain solutes can pass. Some porins are therefore highly selective: *maltoporin*, for example, preferentially allows maltose and maltose oligomers to cross the outer membrane of *E. coli*.

The *FepA protein* is a more complex example of a β barrel transport protein (example 4 in Figure 10–26). It transports iron ions across the bacterial outer membrane. It is constructed from 22 β strands, and a large globular domain completely fills the inside of the barrel. Iron ions bind to this domain, which is thought to undergo a large conformational change to transfer the iron across the membrane.

Not all β barrel proteins are transport proteins. Some form smaller barrels that are completely filled by amino acid side chains that project into the center of the barrel. These proteins function as receptors or enzymes (examples 1 and 2 in Figure 10–26), and the barrel serves as a rigid anchor, which holds the protein in the membrane and orients the cytosolic loops that form binding sites for specific intracellular molecules.

Although β barrel proteins have various functions, they are largely restricted to bacterial, mitochondrial, and chloroplast outer membranes. Most multipass transmembrane proteins in eucaryotic cells and in the bacterial plasma membrane are constructed from transmembrane α helices. The helices can slide against each other, allowing conformational changes in the protein that can open and shut ion channels, transport solutes, or transduce extracellular signals into intracellular ones. In β barrel proteins, by contrast, hydrogen bonds bind each β strand rigidly to its neighbors, making conformational changes within the wall of the barrel unlikely.

Many Membrane Proteins Are Glycosylated

Most transmembrane proteins in animal cells are glycosylated. As in glycolipids, the sugar residues are added in the lumen of the ER and the Golgi apparatus

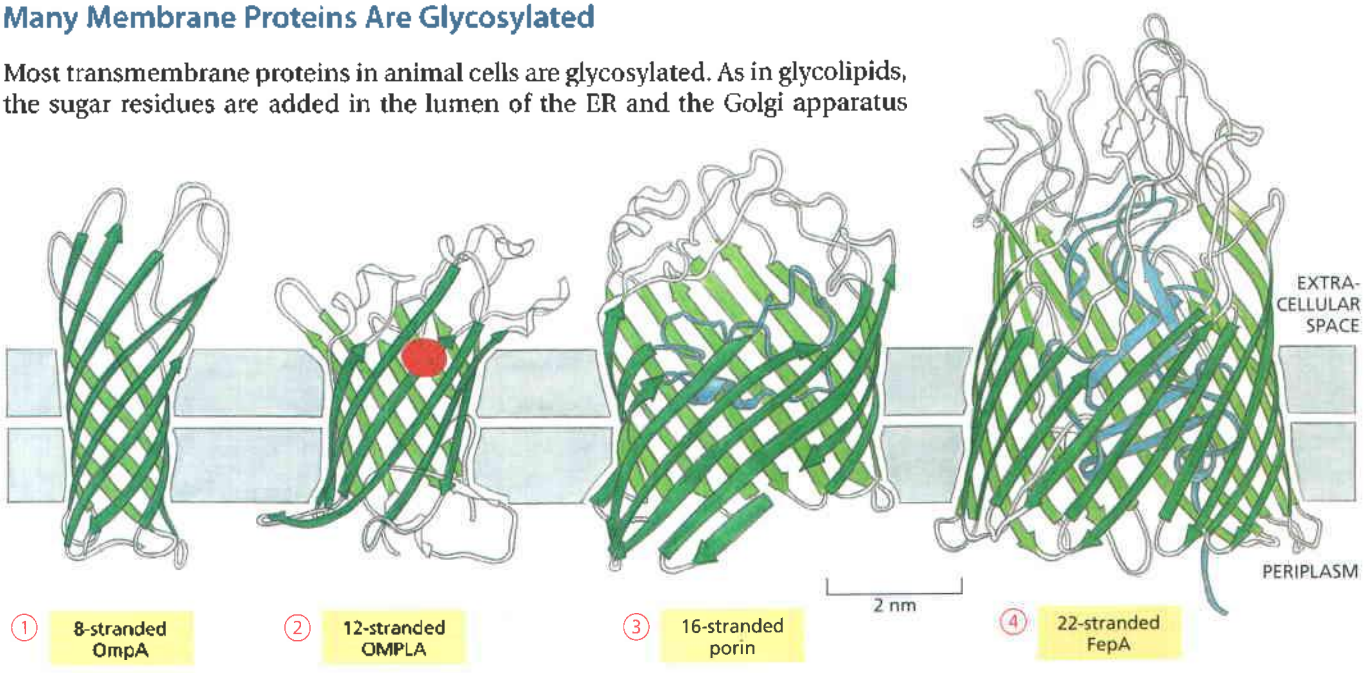


Figure 10–26 β barrels formed from different numbers of β strands. <TGCT> (1) The *E. coli* OmpA protein serves as a receptor for a bacterial virus. (2) The *E. coli* OMPLA protein is an enzyme (a lipase) that hydrolyzes lipid molecules. The amino acids that catalyze the enzymatic reaction (shown in *red*) protrude from the outside surface of the barrel. (3) A porin from the bacterium *Rhodobacter capsulatus* forms a water-filled pore across the outer membrane. The diameter of the channel is restricted by loops (shown in *blue*) that protrude into the channel. (4) The *E. coli* FepA protein transports iron ions. The inside of the barrel is completely filled by a globular protein domain (shown in *blue*) that contains an iron-binding site (not shown). This domain is thought to change its conformation to transport the bound iron, but the molecular details of the changes are not known.

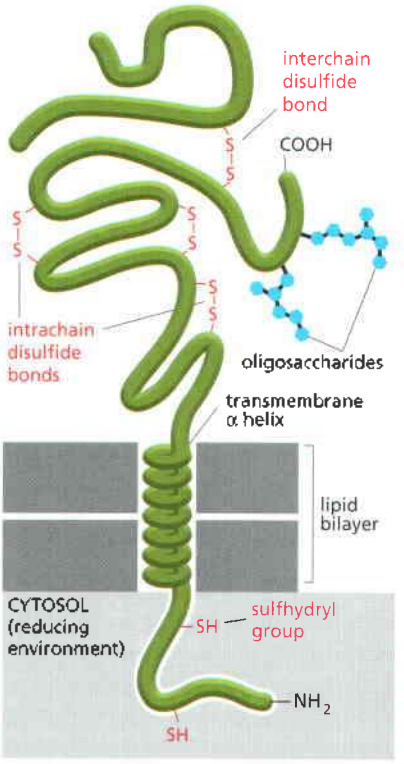


Figure 10–27 A single-pass transmembrane protein. Note that the polypeptide chain traverses the lipid bilayer as a right-handed α helix and that the oligosaccharide chains and disulfide bonds are all on the noncytosolic surface of the membrane. The sulfhydryl groups in the cytosolic domain of the protein do not normally form disulfide bonds because the reducing environment in the cytosol maintains these groups in their reduced (–SH) form.

(discussed in Chapters 12 and 13). For this reason, the oligosaccharide chains are always present on the noncytosolic side of the membrane. Another important difference between proteins (or parts of proteins) on the two sides of the membrane results from the reducing environment of the cytosol. This environment decreases the likelihood that intrachain or interchain disulfide (S–S) bonds will form between cysteines on the cytosolic side of membranes. These bonds form on the noncytosolic side, where they can help stabilize either the folded structure of the polypeptide chain or its association with other polypeptide chains (**Figure 10–27**).

Because most plasma membrane proteins are glycosylated, carbohydrates extensively coat the surface of all eucaryotic cells. These carbohydrates occur as oligosaccharide chains covalently bound to membrane proteins (glycoproteins) and lipids (glycolipids). They also occur as the polysaccharide chains of integral membrane *proteoglycan* molecules. Proteoglycans, which consist of long polysaccharide chains linked covalently to a protein core, are found mainly outside the cell, as part of the extracellular matrix (discussed in Chapter 19). But, for some proteoglycans, the protein core either extends across the lipid bilayer or is attached to the bilayer by a glycosylphosphatidylinositol (GPI) anchor.

The terms cell coat or glycocalyx are sometimes used to describe the carbohydrate-rich zone on the cell surface. This **carbohydrate layer** can be visualized by various stains, such as ruthenium red (**Figure 10–28A**), as well as by its affinity for carbohydrate-binding proteins called **lectins**, which can be labeled with a fluorescent dye or some other visible marker. Although most of the sugar groups are attached to intrinsic plasma membrane molecules, the carbohydrate layer also contains both glycoproteins and proteoglycans that have been secreted into the extracellular space and then adsorbed onto the cell surface (Figure 10–28B). Many of these adsorbed macromolecules are components of the extracellular matrix, so that the boundary between the plasma membrane and the extracellular matrix is often not sharply defined. One of the many functions of the carbohydrate layer is to protect cells against mechanical and chemical damage; it also keeps various other cells at a distance, preventing unwanted protein–protein interactions.

The oligosaccharide side chains of glycoproteins and glycolipids are enormously diverse in their arrangement of sugars. Although they usually contain fewer than 15 sugars, they are often branched, and the sugars can be bonded together by various covalent linkages—unlike the amino acids in a polypeptide chain, which are all linked by identical peptide bonds. Even three sugars can be put together to form hundreds of different trisaccharides. Both the diversity and the exposed position of the oligosaccharides on the cell surface make them especially well suited to function in specific cell-recognition processes. As we discuss in Chapter 19, plasma membrane-bound lectins that recognize specific oligosaccharides on cell-surface glycolipids and glycoproteins mediate a variety of transient cell–cell adhesion processes, including those occurring in sperm–egg interactions, blood clotting, lymphocyte recirculation, and inflammatory responses.

Membrane Proteins Can Be Solubilized and Purified in Detergents

In general, only agents that disrupt hydrophobic associations and destroy the lipid bilayer can solubilize transmembrane proteins (and some other tightly bound membrane proteins). The most useful of these for the membrane biochemist are **detergents**, which are small amphiphilic molecules of variable

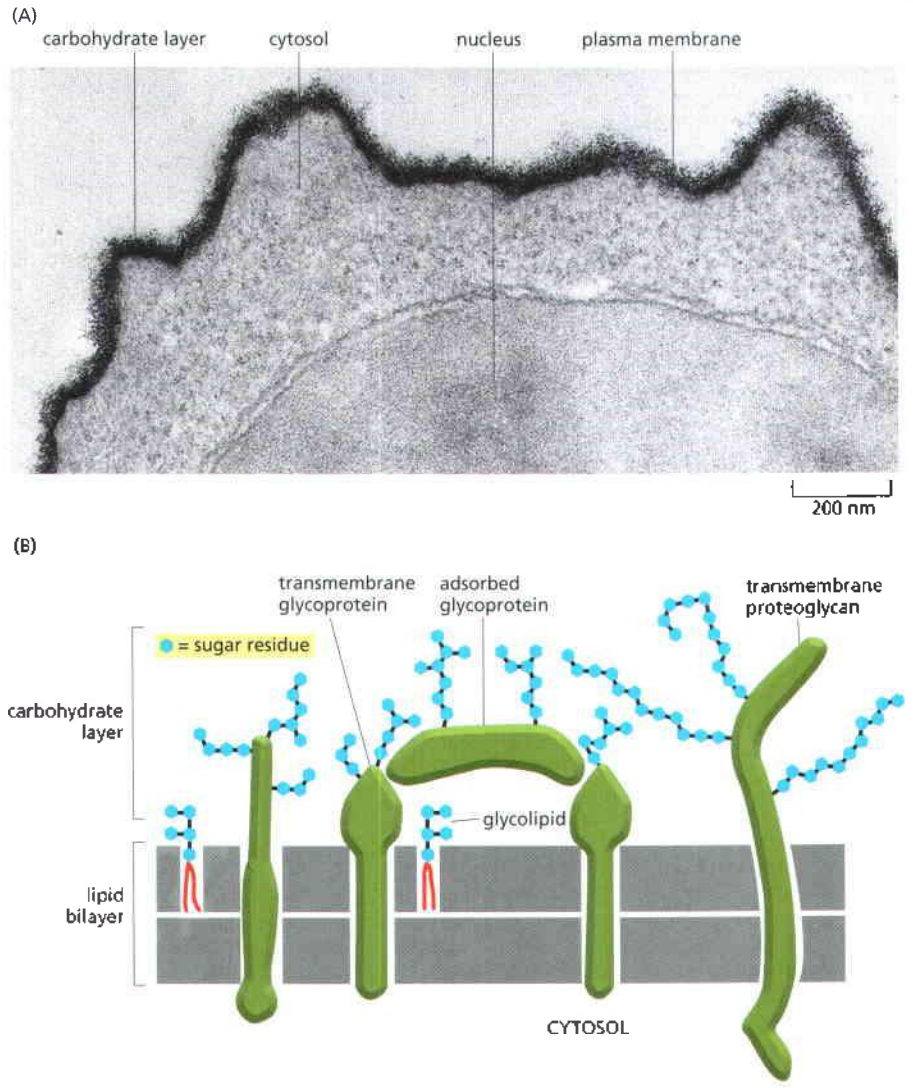


Figure 10–28 The carbohydrate layer on the cell surface. This electron micrograph of the surface of a lymphocyte stained with ruthenium red emphasizes the thick carbohydrate-rich layer surrounding the cell. (B) The carbohydrate layer is made up of the oligosaccharide side chains of glycolipids and integral membrane glycoproteins and the polysaccharide chains on integral membrane proteoglycans. In addition, adsorbed glycoproteins, and adsorbed proteoglycans (not shown) contribute to the carbohydrate layer in many cells. Note that all of the carbohydrate is on the noncytosolic surface of the membrane. (A, courtesy of Audrey M. Glauert and G.M.W. Cook.)

structure. Detergents are much more soluble in water than lipids. Their polar (hydrophilic) ends can be either charged (ionic), as in *sodium dodecyl sulfate (SDS)*, or uncharged (nonionic), as in *octylglucoside* and Triton (**Figure 10–29**A). At low concentration, detergents are monomeric in solution, but when their concentration is increased above a threshold, called the *critical micelle concentration* or *CMC*, they aggregate to form micelles (Figure 10–29B–C). Detergent molecules rapidly diffuse in and out of micelles, keeping the concentration of monomer in the solution constant, no matter how many micelles are present. Both the CMC and the average number of detergent molecules in a micelle are characteristic properties of each detergent, but they also depend on the temperature, pH, and salt concentration. Detergent solutions are therefore complex systems and are difficult to study.

When mixed with membranes, the hydrophobic ends of detergents bind to the hydrophobic regions of the membrane proteins, where they displace lipid molecules with a collar of detergent molecules. Since the other end of the detergent molecule is polar, this binding tends to bring the membrane proteins into solution as detergent–protein complexes (**Figure 10–30**). Usually, some lipid molecules also remain attached to the protein.

Strong ionic detergents, such as SDS, can solubilize even the most hydrophobic membrane proteins. This allows the proteins to be analyzed by *SDS polyacrylamide-gel electrophoresis* (discussed in Chapter 8), a procedure that has revolutionized the study of membrane proteins. Such strong detergents unfold (denature) proteins by binding to their internal "hydrophobic cores,"

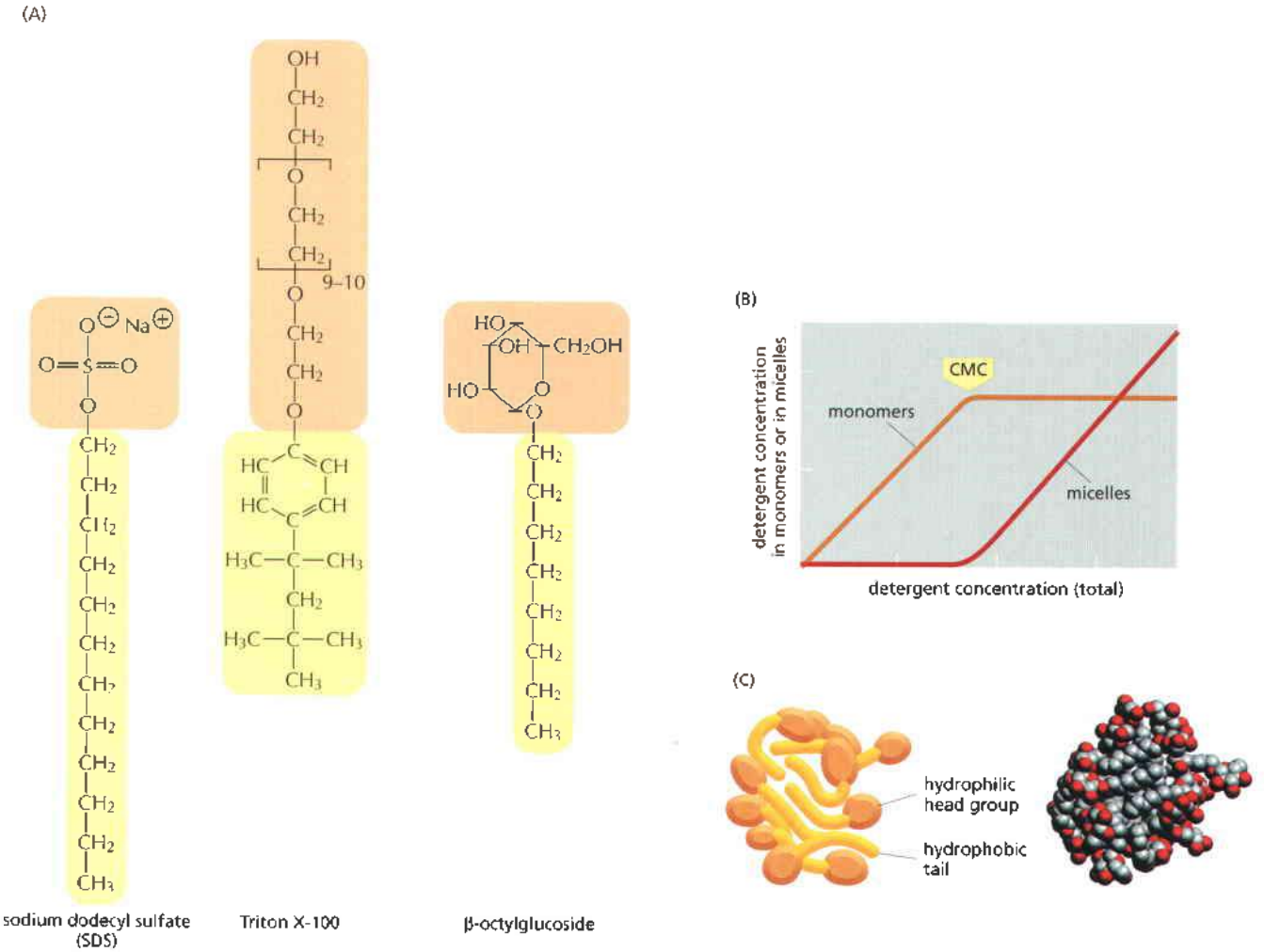


Figure 10–29 The structure and function of detergent micelles. <GCGA> (A) Three commonly used detergents are sodium dodecyl sulfate (SDS), an anionic detergent, and Triton X-100 and β-octylglucoside, two nonionic detergents. Triton X-100 is a mixture of compounds in which the region in brackets is repeated between 9 and 10 times. The hydrophobic portion of each detergent is shown in *yellow*, and the hydrophilic portion is shown in *orange*. (B) At low concentration, detergent molecules are monomeric in solution. As their concentration is increased beyond the critical micelle concentration (CMC), some of the detergent molecules form micelles. Note that the concentration of detergent monomer stays constant above the CMC. (C) Because they have both polar and nonpolar ends, detergent molecules are amphiphilic; and because they are cone-shaped, they form micelles rather than bilayers (see Figure 10–7). Detergent micelles have irregular shapes, and, due to packing constraints, the hydrophobic tails are partially exposed to water. The space-filling model shows the structure of a micelle composed of 20 β-octylglucoside molecules, predicted by molecular dynamics calculations (B, adapted from G. Gunnarsson, B. Jönsson and H. Wennerström, *J. Phys. Chem.* 84:3114–3121, 1980; C, from S. Bogusz, R.M. Venable and R.W. Pastor, *J. Phys. Chem. B.* 104:5462–5470, 2000. With permission from the American Chemical Society.)

thereby rendering the proteins inactive and unusable for functional studies. Nonetheless, proteins can be readily separated and purified in their SDS-denatured form. In some cases, removal of the detergent allows the purified protein to renature, with recovery of functional activity.

Many hydrophobic membrane proteins can be solubilized and then purified in an active form by the use of mild detergents. These detergents cover the hydrophobic regions on membrane-spanning segments that become exposed after lipid removal but do not unfold the protein. If the detergent concentration of a solution of solubilized membrane proteins is reduced (by dilution, for example), membrane proteins do not remain soluble. In the presence of an excess of phospholipid molecules in such a solution, membrane proteins incorporate into small liposomes that form spontaneously. In this way, functionally active membrane protein systems can be reconstituted from purified components, providing a powerful means of analyzing the activities of membrane transporters, ion channels, signaling receptors, and so on (**Figure 10–31**). Such

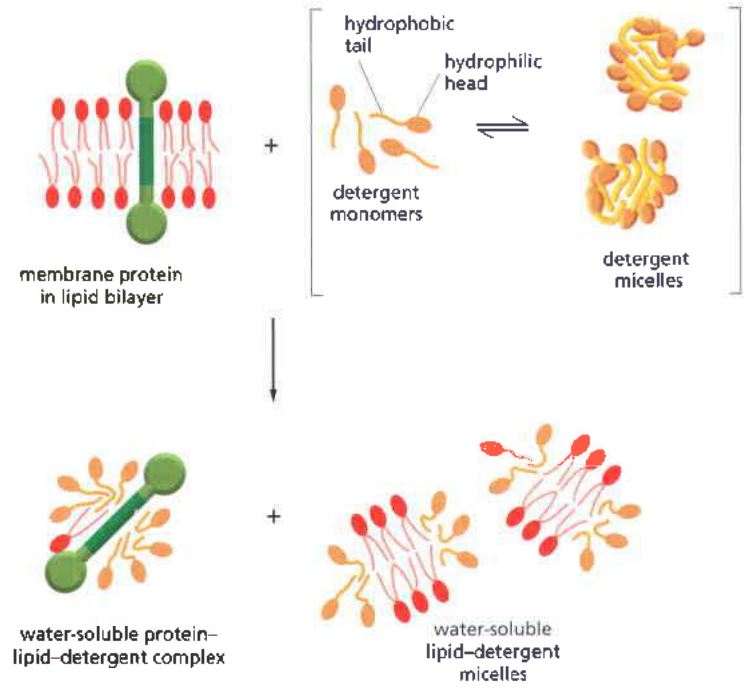


Figure 10–30 Solubilizing membrane proteins with a mild nonionic detergent. The detergent disrupts the lipid bilayer and brings the proteins into solution as protein–lipid– detergent complexes. The phospholipids in the membrane are also solubilized by the detergent.

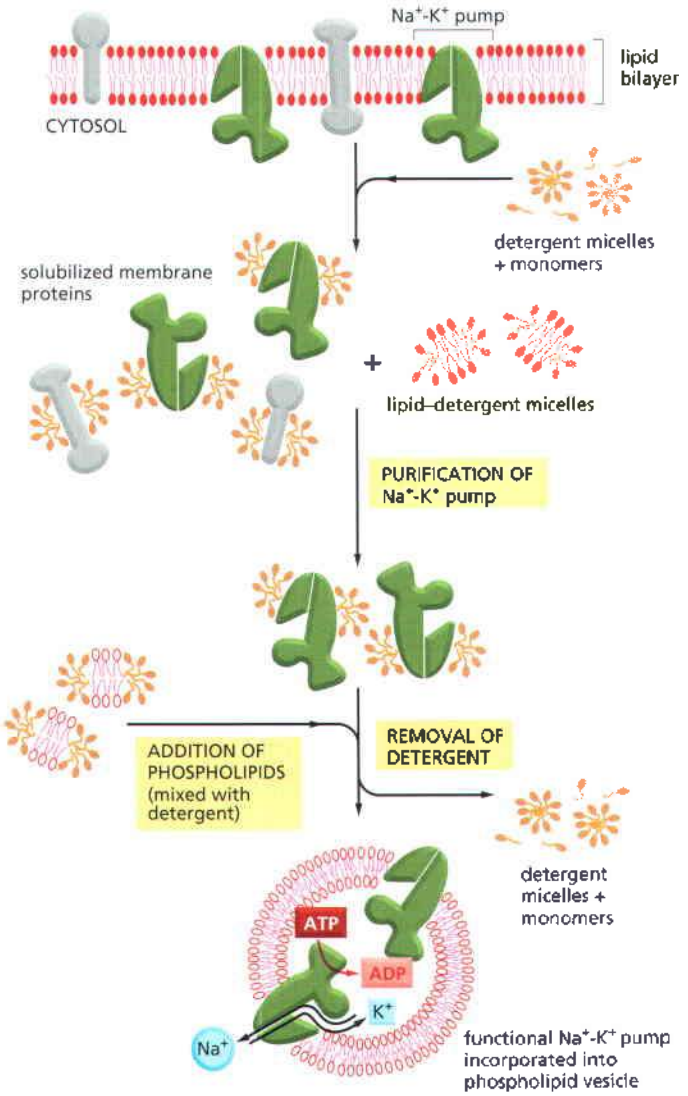


Figure 10–31 The use of mild nonionic detergents for solubilizing, purifying, and reconstituting functional membrane protein systems. In this example, functional Na^+-K^+ pump molecules are purified and incorporated into phospholipid vesicles. The Na^+-K^+ pump is an ion pump that is present in the plasma membrane of most animal cells; it uses the energy of ATP hydrolysis to pump Na^+ out of the cell and K^+ in, as discussed in Chapter 11.

functional reconstitution, for example, provided proof for the hypothesis that the transmembrane ATPases use H^+ gradients in mitochondrial, chloroplast, and bacterial membranes to synthesize ATP.

Detergents have also played a crucial part in the purification and crystallization of membrane proteins. The development of new detergents and new expression systems producing large quantities of membrane proteins from cDNA clones has led to a rapid increase in the number of structures of membrane proteins and protein complexes that are known.

Bacteriorhodopsin Is a Light-Driven Proton Pump That Traverses the Lipid Bilayer as Seven α Helices

In Chapter 11, we consider how multipass transmembrane proteins mediate the selective transport of small hydrophilic molecules across cell membranes. But a detailed understanding of how a membrane transport protein actually works requires precise information about its three-dimensional structure in the bilayer. *Bacteriorhodopsin* was the first membrane transport protein whose structure was determined. It has remained the prototype of many multipass membrane proteins with a similar structure, and it merits a brief digression here.

The "purple membrane" of the archaean *Halobacterium salinarum* is a specialized patch in the plasma membrane that contains a single species of protein molecule, **bacteriorhodopsin** (Figure 10–32). Each bacteriorhodopsin molecule contains a single light-absorbing group, or chromophore (called *retinal*), which gives the protein its purple color. Retinal is vitamin A in its aldehyde form and is identical to the chromophore found in *rhodopsin* of the photoreceptor cells of the vertebrate eye (discussed in Chapter 15). Retinal is covalently linked to a lysine side chain of the bacteriorhodopsin protein. When activated by a single photon of light, the excited chromophore changes its shape and causes a series of small conformational changes in the protein, resulting in the transfer of one H^+ from the inside to the outside of the cell (Figure 10–33). In bright light, each bacteriorhodopsin molecule can pump several hundred protons per second. The light-driven proton transfer establishes an H^+ gradient across the plasma membrane, which in turn drives the production of ATP by a second protein in the cell's plasma membrane. The energy stored in the H^+ gradient also drives other energy-requiring processes in the cell. Thus, bacteriorhodopsin converts solar energy into a proton gradient, which provides energy to the archaeal cell.

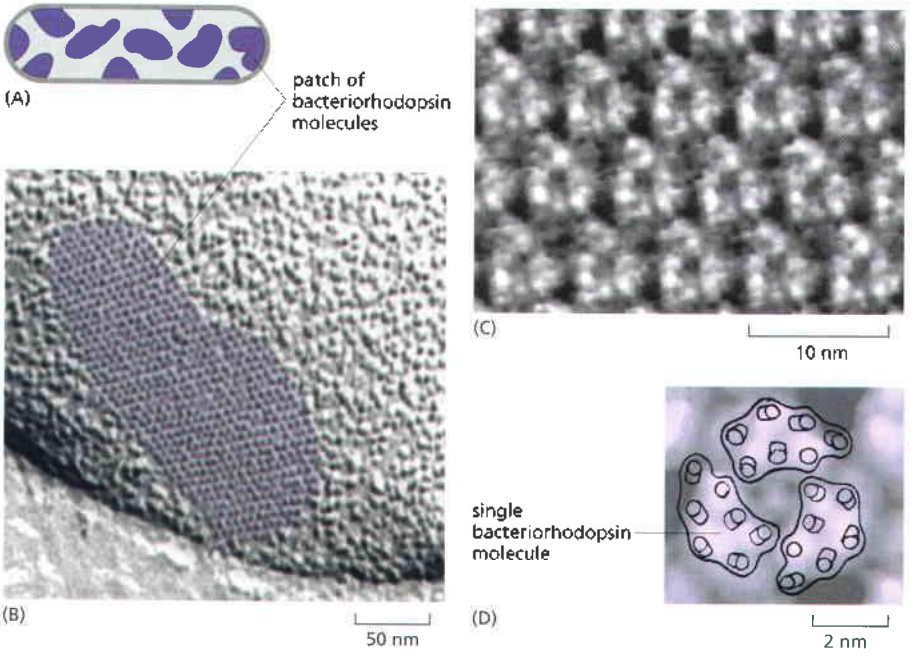
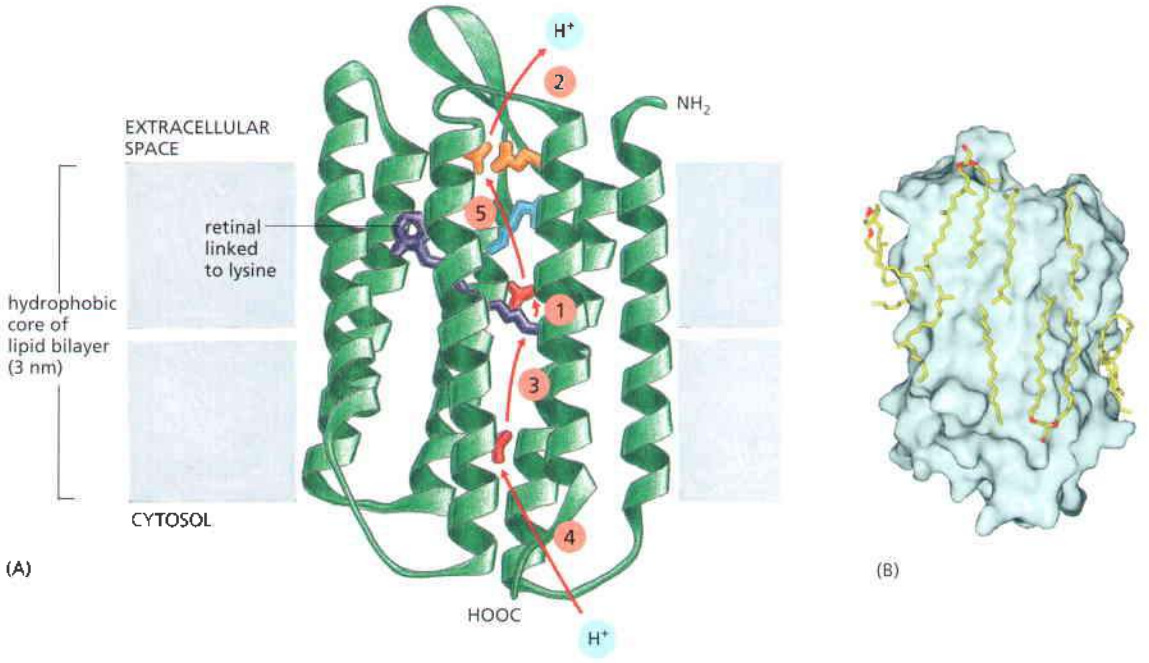


Figure 10–32 Patches of purple membrane, which contain bacteriorhodopsin in the archaean *Halobacterium salinarum*. (A) These archaea live in saltwater pools, where they are exposed to sunlight. They have evolved a variety of light-activated proteins, including bacteriorhodopsin, which is a light-activated proton pump in the plasma membrane. (B) The bacteriorhodopsin molecules in the purple membrane patches are tightly packed into two-dimensional crystalline arrays. (C) Details of the molecular surface visualized by atomic force microscopy. With this technique individual bacteriorhodopsin molecules can be seen. (D) Outline of the approximate locations of three bacteriorhodopsin monomers and their individual α helices in the images shown in (B). (B–D, courtesy of Dieter Oesterhelt.)



The numerous bacteriorhodopsin molecules in the purple membrane are arranged as a planar two-dimensional crystal. The regular packing has made it possible to determine the three-dimensional structure and orientation of bacteriorhodopsin in the membrane to moderate resolution (3 Å) by an approach that uses a combination of electron microscopy and electron diffraction analysis. This procedure, known as **electron crystallography**, is analogous to the study of three-dimensional crystals of soluble proteins by x-ray diffraction analysis. It has provided the first structural views of many membrane proteins that were found to be difficult to crystallize from detergent solutions. For bacteriorhodopsin, the structure obtained by electron crystallography was later confirmed and extended to very high resolution by x-ray crystallography. Each bacteriorhodopsin molecule is folded into seven closely packed α helices (each containing about 25 amino acids), which pass through the lipid bilayer at slightly different angles. By obtaining very well ordered protein crystals and freezing them at very low temperatures, it has also been possible to determine the structures of some of the protein's intermediate conformations during its H^+ pumping cycle.

Bacteriorhodopsin is a member of a large superfamily of membrane proteins with similar structures but different functions. For example, rhodopsin in rod cells of the vertebrate retina and many cell-surface receptor proteins that bind extracellular signal molecules are also built from seven transmembrane α helices. These proteins function as signal transducers rather than as transporters: each responds to an extracellular signal by activating a GTP-binding protein (G protein) inside the cell and are therefore called *G-protein-coupled receptors* (*GPCRs*), as we discuss in Chapter 15. Although the structures of bacteriorhodopsins and GPCRs are strikingly similar, they show no sequence similarity and thus probably belong to two evolutionarily distant branches of an ancient protein family.

The high-resolution crystal structure of bacteriorhodopsin reveals many lipid molecules that are bound in specific places on the protein surface (Figure 10–33B). Interactions with specific lipids are thought to help stabilize many membrane proteins, which work best and crystallize more readily if some of the lipids remain bound during detergent extraction, or if specific lipids are added back to the proteins in detergent solutions. The specificity of these lipid–protein interactions helps explain why eucaryotic membranes contain such a variety of lipids, with head groups that differ in size, shape, and charge. We can think of the membrane lipids as constituting a two-dimensional solvent for the proteins in the membrane, just as water constitutes a three-dimensional solvent for proteins

Figure 10–33 The three-dimensional structure of a bacteriorhodopsin molecule. <TTAA> (A) The polypeptide chain crosses the lipid bilayer seven times as α helices. The location of the retinal chromophore *(purple)* and the probable pathway taken by protons during the light-activated pumping cycle are shown. The first and key step is the passing of a H^+ from the chromophore to the side chain of aspartic acid 85 *(red,* located next to the chromophore) that occurs upon absorption of a photon by the chromophore. Subsequently, other H^+ transfers—in the numerical order indicated and utilizing the hydrophilic amino acid side chains that line a path through the membrane—complete the pumping cycle and return the enzyme to its starting state. Color code: glutamic acid *(orange)*, aspartic acid *(red)*, arginine *(blue)*. (B) The high-resolution crystal structure of bacteriorhodopsin shows many lipid molecules (*yellow* with *red* head groups) that are tightly bound to specific places on the surface of the protein. (A, adapted from H. Luecke et al., *Science* 286:255–260, 1999. With permission from AAAS; B, from H. Luecke et al., *J. Mol. Biol.* 291:899–911, 1999. With permission from Academic Press.)

in an aqueous solution. Some membrane proteins can function only in the presence of specific lipid head groups, just as many enzymes in aqueous solution require a particular ion for activity.

Membrane Proteins Often Function as Large Complexes

Many membrane proteins function as part of multicomponent complexes, several of which have been studied by x-ray crystallography. One is a bacterial *photosynthetic reaction center*, which was the first transmembrane protein complex to be crystallized and analyzed by x-ray diffraction. The results of this analysis were of general importance to membrane biology because they showed for the first time how multiple polypeptides associate in a membrane to form a complex protein machine (**Figure 10–34**). In Chapter 14, we discuss how such photosynthetic complexes function to capture light energy and use it to pump protons across the membrane. Many of the membrane protein complexes involved in photosynthesis, proton pumping, and electron transport are even larger than the photosynthetic reaction center. The enormous photosystem II complex from cyanobacteria, for example, contains 19 protein subunits and well over 60 transmembrane helices. Membrane proteins are often arranged in large complexes, not only for harvesting various forms of energy, but also for transducing extracellular signals into intracellular ones (discussed in Chapter 15).

Many Membrane Proteins Diffuse in the Plane of the Membrane

Like most membrane lipids, membrane proteins do not tumble *(flip-flop)* across the lipid bilayer, but they do rotate about an axis perpendicular to the plane of the bilayer *(rotational diffusion)*. In addition, many membrane proteins are able

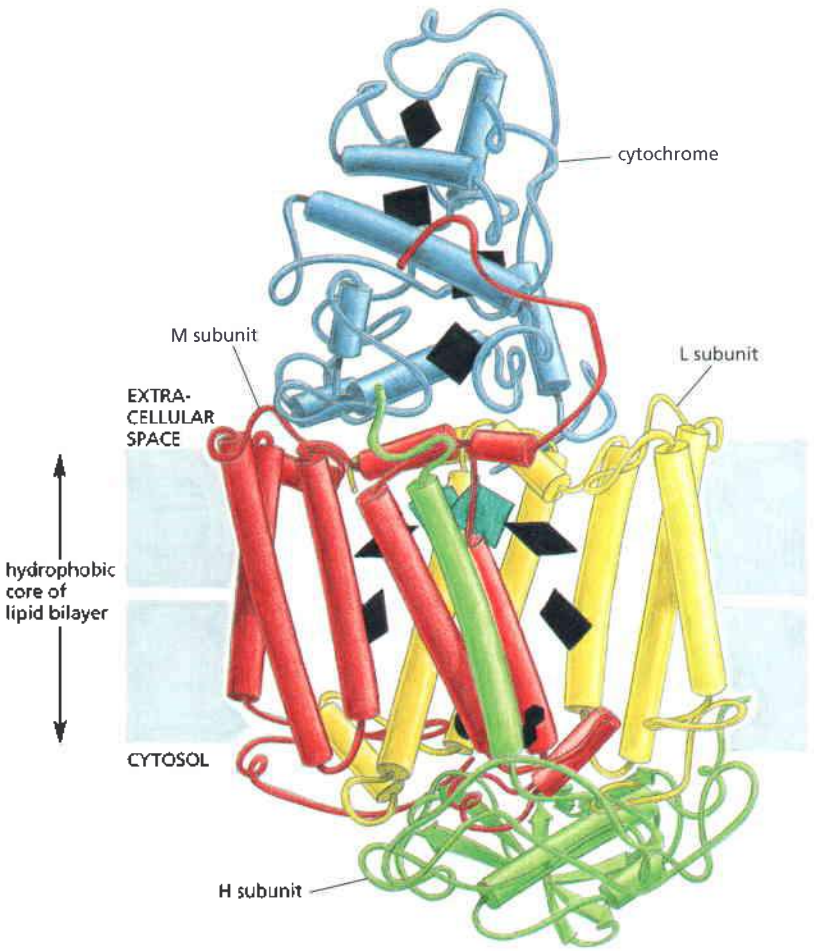


Figure 10–34 The three-dimensional structure of the photosynthetic reaction center of the bacterium *Rhodopseudomonas viridis*. <ATCA> The structure was determined by x-ray diffraction analysis of crystals of this transmembrane protein complex. The complex consists of four subunits L, M, H, and a cytochrome. The L and M subunits form the core of the reaction center, and each contains five α helices that span the lipid bilayer. The locations of the various electron carrier coenzymes are shown in *black*. Note that the coenzymes are arranged in the spaces between the helices. The *special pair* of chlorophyll molecules (discussed in Chapter 14) is shown in *turquoise*. (Adapted from a drawing by J. Richardson based on data from J. Deisenhofer et al., *Nature* 318:618–624, 1985. With permission from Macmillan Publishers Ltd.)

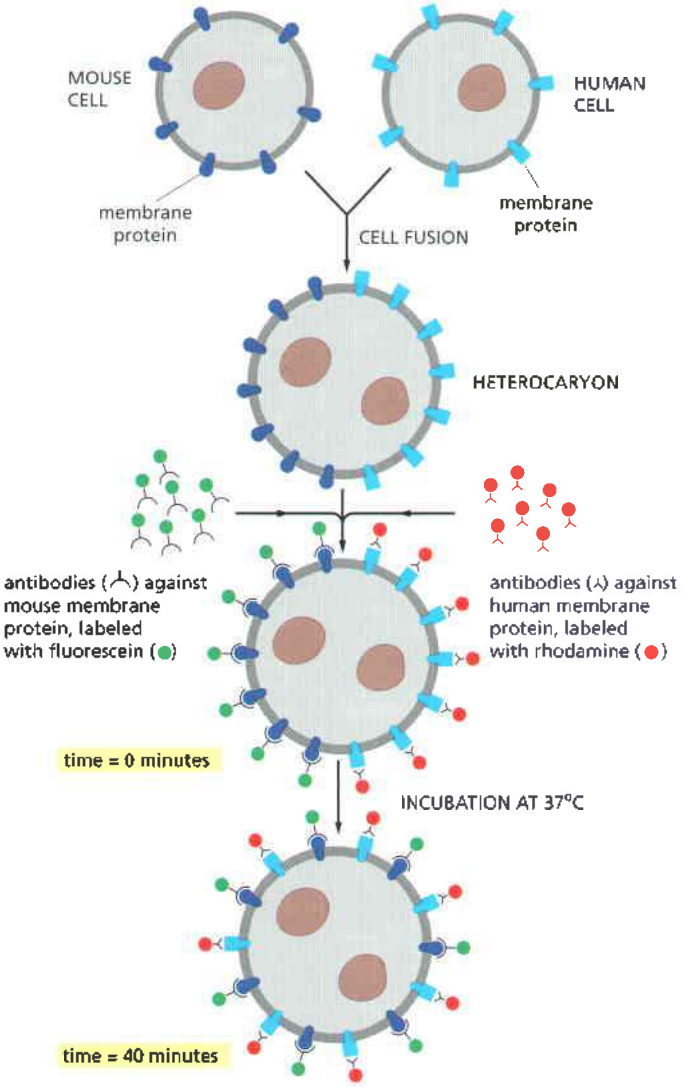


Figure 10–35 An experiment demonstrating the diffusion of proteins in the plasma membrane of mouse–human hybrid cells. The mouse and human proteins are initially confined to their own halves of the newly formed heterocaryon plasma membrane, but they intermix over time. The two antibodies used to visualize the proteins can be distinguished in a fluorescence microscope because fluorescein is green and rhodamine is red. (Based on L.D. Frye and M. Edidin, *J. Cell Sci.* 7:319–335, 1970. With permission from The Company of Biologists.)

to move laterally within the membrane *(lateral diffusion)*. An experiment in which mouse cells were artificially fused with human cells to produce hybrid cells *(heterocaryons)* provided the first direct evidence that some plasma membrane proteins are mobile in the plane of the membrane. Two differently labeled antibodies were used to distinguish selected mouse and human plasma membrane proteins. Although at first the mouse and human proteins were confined to their own halves of the newly formed heterocaryon, the two sets of proteins diffused and mixed over the entire cell surface in about half an hour (**Figure 10–35**).

The lateral diffusion rates of membrane proteins can be measured by using the technique of *fluorescence recovery after photobleaching (FRAP)*. The method usually involves marking the membrane protein of interest with a specific fluorescent group. This can be done either with a fluorescent ligand such as a fluorophore-labeled antibody that binds to the protein or with recombinant DNA technology to express the protein fused to green fluorescent protein (GFP) (discussed in Chapter 9). The fluorescent group is then bleached in a small area of membrane by a laser beam, and the time taken for adjacent membrane proteins carrying unbleached ligand or GFP to diffuse into the bleached area is measured (**Figure 10–36**A). A complementary technique is *fluorescence loss in photobleaching (FLIP)*. Here, a laser beam continuously irradiates a small area of membrane to bleach all the fluorescent molecules that diffuse into it, thereby gradually depleting the surrounding membrane of fluorescently labeled molecules (Figure 10–36B). From such FRAP and FLIP measurements, we can calculate the diffusion coefficient for the marked cell-surface protein. The values

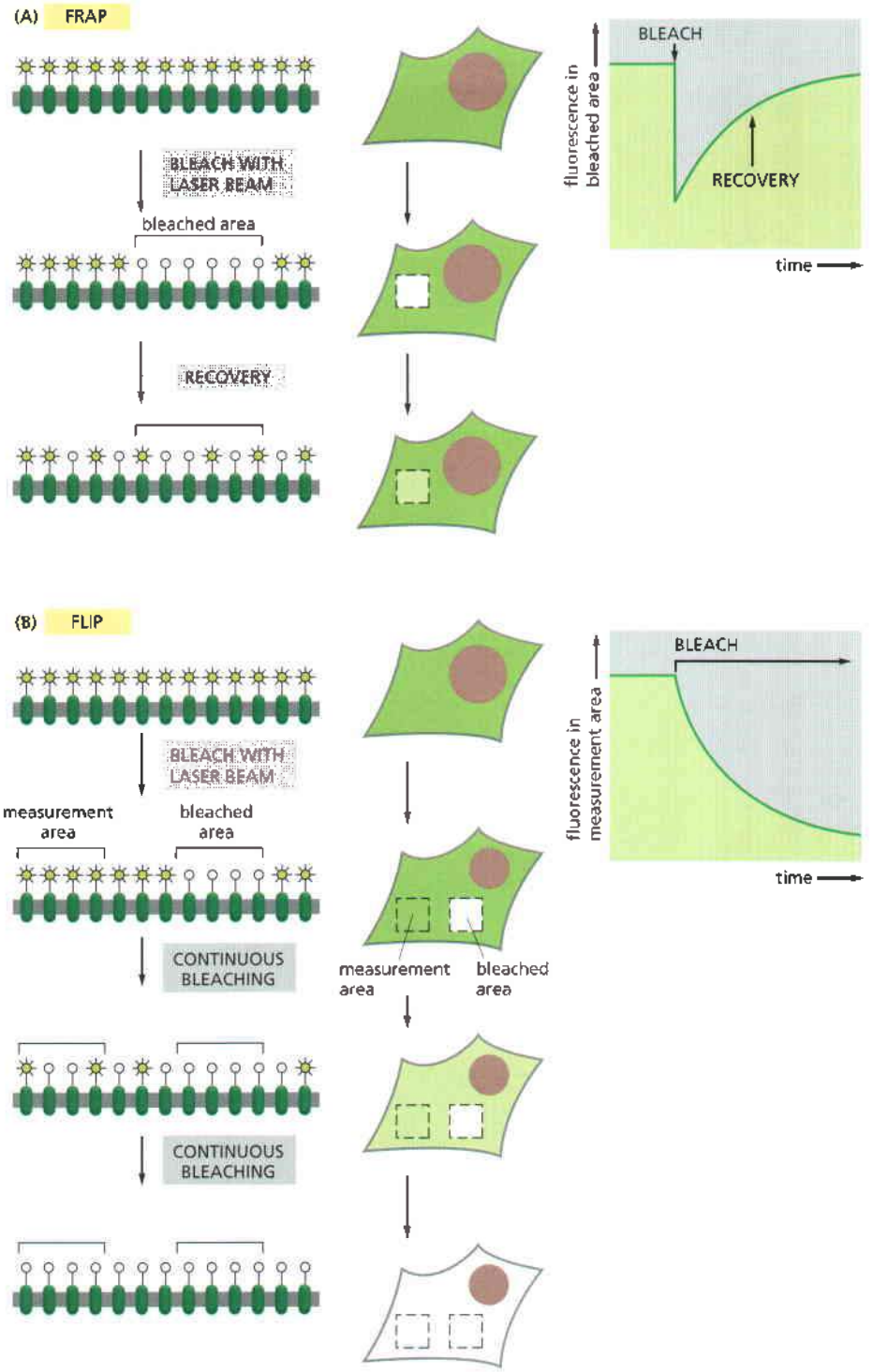


Figure 10–36 Measuring the rate of lateral diffusion of a membrane protein by photobleaching techniques. <ATGT> A specific protein of interest can be expressed as a fusion protein with green fluorescent protein (GFP), which is intrinsically fluorescent. (A) In the FRAP technique, fluorescent molecules are bleached in a small area using a laser beam. The fluorescence intensity recovers as the bleached molecules diffuse away and unbleached molecules diffuse into the irradiated area (shown here in side and top views). The diffusion coefficient is calculated from a graph of the rate of recovery: the greater the diffusion coefficient of the membrane protein, the faster the recovery. (B) In the FLIP technique, a small area of membrane is irradiated continuously, and fluorescence is measured in a separate area. Fluorescence in the second area progressively decreases as fluorescent proteins diffuse out and bleached molecules diffuse in; eventually, all of the fluorescent protein molecules are bleached, as long as they are mobile and not permanently anchored to the cytoskeleton or extracellular matrix.

of the diffusion coefficients for different membrane proteins in different cells are highly variable, because interactions with other proteins impede the diffusion of the proteins to varying degrees. Measurements of proteins that are minimally impeded in this way indicate that cell membranes have a viscosity comparable to that of olive oil.

One drawback to the FRAP and FLIP techniques is that they monitor the movement of large populations of molecules in a relatively large area of membrane; one cannot follow individual protein molecules. If a protein fails to migrate into a bleached area, for example, one cannot tell whether the molecule is truly immobile or just restricted in its movement to a very small region of membrane—perhaps by cytoskeletal proteins. *Single-particle tracking* techniques overcome this problem by labeling individual membrane molecules with antibodies coupled to fluorescent dyes or tiny gold particles and tracking their

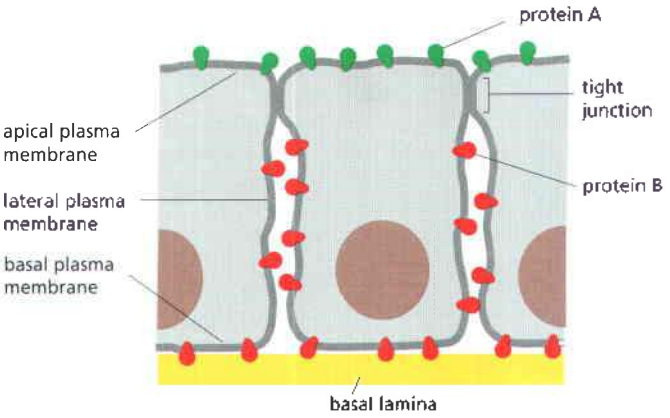


Figure 10–37 How membrane molecules can be restricted to a particular membrane domain. In this drawing of an epithelial cell, protein A (in the apical membrane) and protein B (in the basal and lateral membranes) can diffuse laterally in their own domains but are prevented from entering the other domain, at least partly by the specialized cell junction called a tight junction. Lipid molecules in the outer (noncytosolic) monolayer of the plasma membrane are likewise unable to diffuse between the two domains; lipids in the inner (cytosolic) monolayer, however, are able to do so (not shown). The basal lamina is a thin mat of extracellular matrix that separates epithelial sheets from other tissues (discussed in Chapter 19).

movement by video microscopy. Using single-particle tracking, one can record the diffusion path of a single membrane protein molecule over time. Results from all of these techniques indicate that plasma membrane proteins differ widely in their diffusion characteristics, as we now discuss.

Cells Can Confine Proteins and Lipids to Specific Domains Within a Membrane

The recognition that biological membranes are two-dimensional fluids was a major advance in understanding membrane structure and function. It has become clear, however, that the picture of a membrane as a lipid sea in which all proteins float freely is greatly oversimplified. Many cells confine membrane proteins to specific regions in a continuous lipid bilayer. We have already discussed how bacteriorhodopsin molecules in the purple membrane of *Halobacterium* assemble into large two-dimensional crystals, in which individual protein molecules are relatively fixed in relationship to one another (see Figure 10–32); large aggregates of this kind diffuse very slowly.

In epithelial cells, such as those that line the gut or the tubules of the kidney, certain plasma membrane enzymes and transport proteins are confined to the apical surface of the cells, whereas others are confined to the basal and lateral surfaces (**Figure 10–37**). This asymmetric distribution of membrane proteins is often essential for the function of the epithelium, as we discuss in Chapters 11 and 19. The lipid compositions of these two membrane domains are also different, demonstrating that epithelial cells can prevent the diffusion of lipid as well as protein molecules between the domains. Experiments with labeled lipids, however, suggest that only lipid molecules in the outer monolayer of the membrane are confined in this way. The barriers set up by a specific type of intercellular junction (called a *tight junction*, discussed in Chapter 19) maintain the separation of both protein and lipid molecules. Clearly, the membrane proteins that form these intercellular junctions cannot be allowed to diffuse laterally in the interacting membranes.

A cell can also create membrane domains without using intercellular junctions. The mammalian spermatozoon, for instance, is a single cell that consists of several structurally and functionally distinct parts covered by a continuous plasma membrane. When a sperm cell is examined by immunofluorescence microscopy with a variety of antibodies, each of which react with a specific cell-surface molecule, the plasma membrane is found to consist of at least three distinct domains (**Figure 10–38**). Some of the membrane molecules are able to diffuse freely within the confines of their own domain. The molecular nature of the "fence" that prevents the molecules from leaving their domain is not known. Many other cells have similar membrane fences that confine membrane protein diffusion to certain membrane domains. The plasma membrane of nerve cells, for example, contains a domain enclosing the cell body and dendrites, and another enclosing the axon. In this case, it is thought that a belt of actin filaments tightly associated with the plasma membrane at the cell-body–axon junction forms part of the barrier.

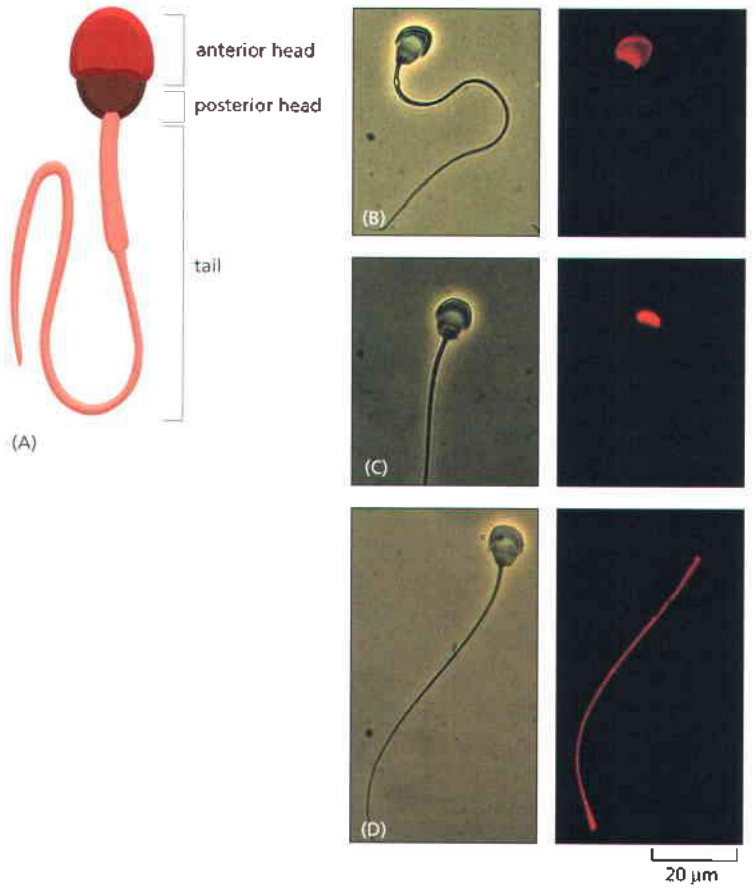


Figure 10–38 Three domains in the plasma membrane of a guinea pig sperm. (A) A drawing of a guinea pig sperm. In the three pairs of micrographs, phase-contrast micrographs are on the *left*, and the same cell is shown with cell-surface immunofluorescence staining on the *right*. Different monoclonal antibodies selectively label cell-surface molecules on (B) the anterior head, (C) the posterior head, and (D) the tail. (Micrographs courtesy of Selena Carroll and Diana Myles.)

Figure 10–39 shows four common ways of immobilizing specific membrane proteins through protein–protein interactions.

The Cortical Cytoskeleton Gives Membranes Mechanical Strength and Restricts Membrane Protein Diffusion

As shown in Figure 10–39B and C, a common way in which a cell restricts the lateral mobility of specific membrane proteins is to tether them to macromolecular assemblies on either side of the membrane. The characteristic biconcave shape of a red blood cell (**Figure 10–40**), for example, results from interactions of its plasma membrane proteins with an underlying *cytoskeleton*, which consists mainly of a meshwork of the filamentous protein **spectrin**. Spectrin is a long, thin, flexible rod about 100 nm in length. Being the principal component of the red cell cytoskeleton, it maintains the structural integrity and shape of the plasma membrane, which is the red cell's only membrane, as the cell has no nucleus or other organelles. The spectrin cytoskeleton is riveted to the membrane through various membrane proteins. The final result is a deformable, net-like meshwork that covers the entire cytosolic surface of the red cell membrane (**Figure 10–41**). This spectrin-based cytoskeleton enables the red cell to withstand the stress on its membrane as it is forced through narrow capillaries. Mice and humans with genetic abnormalities in spectrin are anemic and have red cells that are spherical (instead of concave) and fragile; the severity of the anemia increases with the degree of spectrin deficiency.

An analogous but much more elaborate and complicated cytoskeletal network exists beneath the plasma membrane of most other cells in our body. This network, which constitutes the cortical region (or **cortex**) of the cytoplasm, is rich in actin filaments, which are attached to the plasma membrane in numerous ways. The cortex of nucleated cells contains proteins that are structurally homologous to spectrin and the other components of the red cell cytoskeleton. We discuss the cortical cytoskeleton in nucleated cells and its interactions with the plasma membrane in Chapter 16.

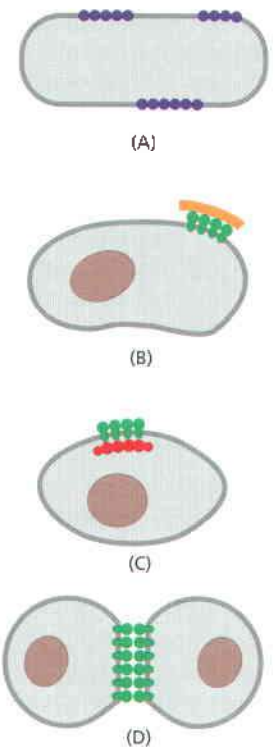


Figure 10–39 Four ways of restricting the lateral mobility of specific plasma membrane proteins. (A) The proteins can self-assemble into large aggregates (as seen for bacteriorhodopsin in the purple membrane of *Halobacterium*); they can be tethered by interactions with assemblies of macromolecules (B) outside or (C) inside the cell; or they can interact with proteins on the surface of another cell (D).

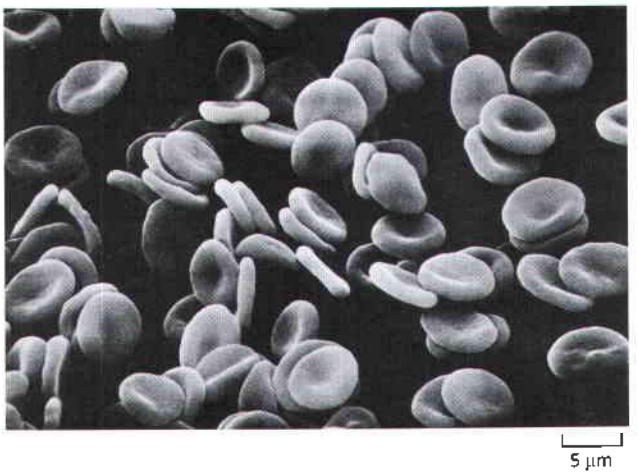


Figure 10–40 A scanning electron micrograph of human red blood cells. <GTAC> The cells have a biconcave shape and lack a nucleus and other organelles. (Courtesy of Bernadette Chailley.)

The cortical cytoskeletal network underlying the plasma membrane restricts diffusion of not only the proteins that are directly anchored to it. Because the *cytoskeletal filaments* are often closely apposed to the cytosolic membrane surface, they can form mechanical barriers that obstruct the free diffusion of membrane proteins. These barriers partition the membrane into small domains, or *corrals* (**Figure 10–42**), which can be either permanent, as in the sperm (see Figure 10–38), or transient. The barriers can be detected when the diffusion of individual membrane proteins is followed by high-speed, single-particle tracking. The proteins diffuse rapidly but are confined within an individual corral; occasionally,

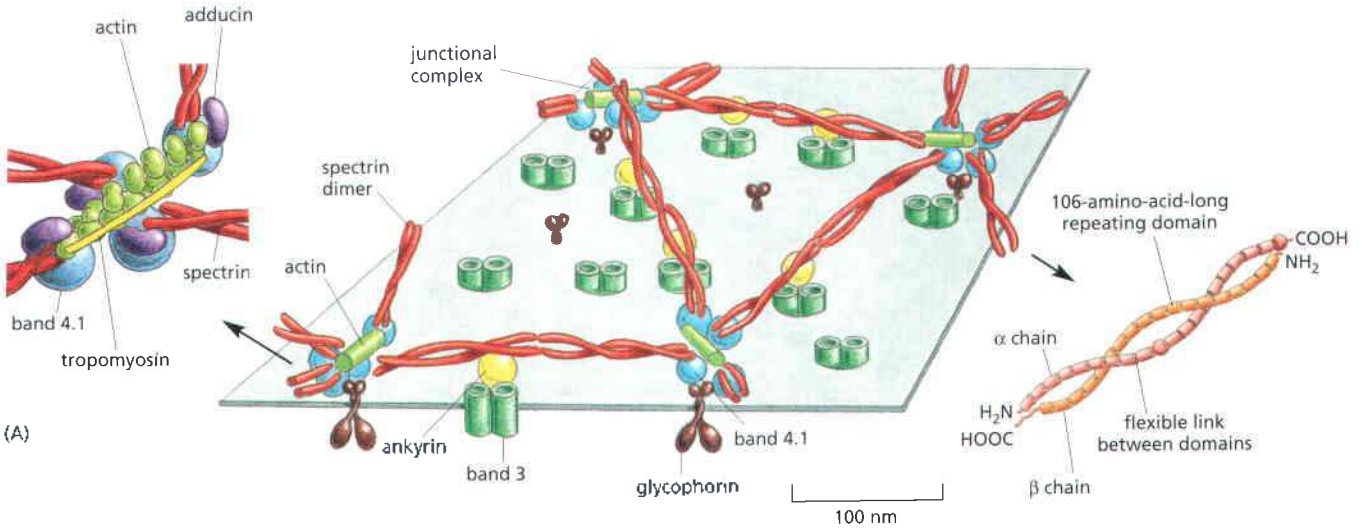
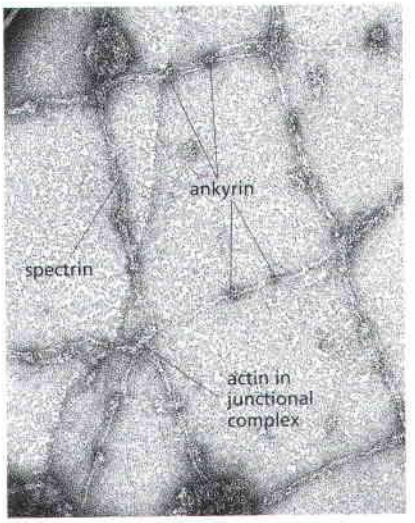


Figure 10–41 The spectrin-based cytoskeleton on the cytosolic side of the human red blood cell plasma membrane. (A) The arrangement shown in the drawing has been deduced mainly from studies on the interactions of purified proteins *in vitro*. Spectrin dimers (enlarged in the box on the *right*) are linked together into a netlike meshwork by "junctional complexes" (enlarged in the box on the *left*). Each spectrin heterodimer consists of two antiparallel, loosely intertwined, flexible polypeptide chains called α and β. The two chains are attached noncovalently to each other at multiple points, including at both ends. Both the α and β chains are composed largely of repeating domains. The junctional complexes are composed of short actin filaments (containing 13 actin monomers), *band 4.1*, *adducin*, and a *tropomyosin* molecule that probably determines the length of the actin filaments. The cytoskeleton is linked to the membrane through two transmembrane proteins—a multipass protein called *band 3* and a single-pass protein called *glycophorin*. The spectrin tetramers bind to some band 3 proteins via *ankyrin* molecules, and to glycophorin and band 3 (not shown) via band 4.1 proteins. (B) The electron micrograph shows the cytoskeleton on the cytosolic side of a red blood cell membrane after fixation and negative staining. The spectrin meshwork has been purposely stretched out to allow the details of its structure to be seen. In a normal cell, the meshwork shown would be much more crowded and occupy only about one-tenth of this area. (B, courtesy of T. Byers and D. Branton, *Proc. Natl Acad. Sci. U.S.A.* 82:6153–6157, 1985. With permission from National Academy of Sciences.)



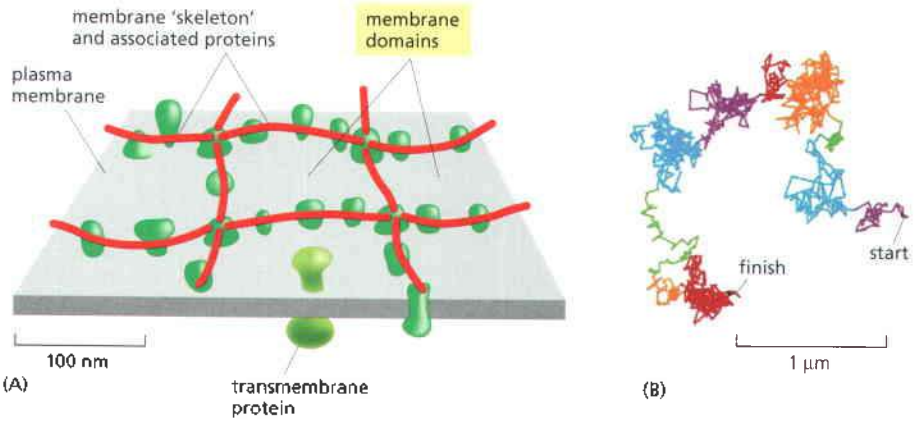


Figure 10–42 Corraling of membrane proteins by cortical cytoskeletal filaments. (A) How cytoskeletal filaments are thought to provide diffusion barriers that divide the membrane into small domains, or corrals. (B) High-speed, single-particle tracking was used to follow the paths of a fluorescently-labeled membrane protein over time. The trace shows that membrane proteins diffuse within a tightly delimited membrane domain (shown by different colors of the trace) and only infrequently escape into a neighboring domain. (Adapted from A. Kusumi et al., *Annu. Rev. Biophys. Biomol. Struct.* 34:351–378, 2005. With permission from Annual Reviews.)

however, thermal motions cause a few cortical filaments to detach transiently from the membrane, allowing the protein to escape into an adjacent corral.

The extent to which a transmembrane protein is confined within a corral depends on its association with other proteins and the size of its cytoplasmic domain; proteins with a large cytosolic domain will have a harder time passing through barriers. When a cell-surface receptor binds its extracellular signal molecules, for example, large protein complexes build up on the cytosolic domain of the receptor, making it more difficult for the receptor to escape from its corral. It is thought that corraling helps concentrate activated signaling complexes, increasing the speed and efficiency of the signaling process (discussed in Chapter 15).

Summary

Whereas the lipid bilayer determines the basic structure of biological membranes, proteins are responsible for most membrane functions, serving as specific receptors, enzymes, transport proteins, and so on. Many membrane proteins extend across the lipid bilayer. Some of these transmembrane proteins are single-pass proteins, in which the polypeptide chain crosses the bilayer as a single α helix. Others are multipass proteins, in which the polypeptide chain crosses the bilayer multiple times—either as a series of α helices or as a β sheet in the form of a closed barrel. All proteins responsible for the transmembrane transport of ions and other small water-soluble molecules are multipass proteins. Some membrane-associated proteins do not span the bilayer but instead are attached to either side of the membrane. Many of these are bound by noncovalent interactions with transmembrane proteins, but others are bound via covalently attached lipid groups. In the plasma membrane of all eucaryotic cells, most of the proteins exposed on the cell surface and some of the lipid molecules in the outer lipid monolayer have oligosaccharide chains covalently attached to them. Like the lipid molecules in the bilayer, many membrane proteins are able to diffuse rapidly in the plane of the membrane. However, cells have ways of immobilizing specific membrane proteins, as well as ways of confining both membrane protein and lipid molecules to particular domains in a continuous lipid bilayer.

PROBLEMS

Which statements are true? Explain why or why not.

10–1 Although lipid molecules are free to diffuse in the plane of the bilayer, they cannot flip-flop across the bilayer unless enzyme catalysts called phospholipid translocators are present in the membrane.

10–2 Whereas all the carbohydrate in the plasma membrane faces outward on the external surface of the cell, all the carbohydrate on internal membranes faces toward the cytosol.

10–3 Although membrane domains with different protein compositions are well known, there are at present no examples of membrane domains that differ in lipid composition.

Discuss the following problems.

10–4 When a lipid bilayer is torn, why does it not seal itself by forming a "hemi-micelle" cap at the edges, as shown in **Figure Q10–1**?

10–5 Margarine is made from vegetable oil by a chemical process. Do you suppose this process converts saturated fatty acids to unsaturated ones, or vice versa? Explain your answer.

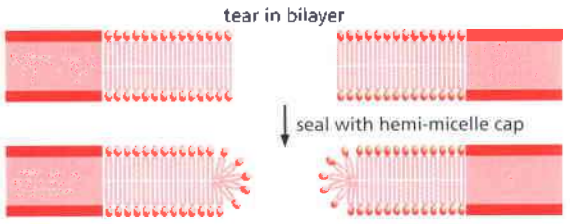


Figure Q10–1 A torn lipid bilayer sealed with a hypothetical "hemi-micelle" cap (Problem 10–4).

10–6 If a lipid raft is typically 70 nm in diameter and each lipid molecule has a diameter of 0.5 nm, about how many lipid molecules would there be in a lipid raft composed entirely of lipid? At a ratio of 50 lipid molecules per protein molecule (50% protein by mass) how many proteins would be in a typical raft? (Neglect the loss of lipid from the raft that would be required to accommodate the protein.)

10–7 A classic paper studied the behavior of lipids in the two monolayers of a membrane by labeling individual molecules with nitroxide groups, which are stable free radicals (**Figure Q10–2**). These spin-labeled lipids can be detected by electron spin-resonance (ESR) spectroscopy, a technique that does not harm living cells. Spin-labeled lipids are introduced into small lipid vesicles, which are then fused with cells, thereby transferring the labeled lipids into the plasma membrane.

The two spin-labeled phospholipids shown in Figure Q10–2 were incorporated into intact human red cell membranes in this way. To determine whether they were introduced equally into the two monolayers of the bilayer, ascorbic acid (vitamin C), which is a water-soluble reducing agent that does not cross membranes, was added to the medium to destroy any nitroxide radicals exposed on the outside of the cell. The ESR signal was followed as a function of time in the presence and absence of ascorbic acid as indicated in **Figure Q10–3A** and B.

A. Ignoring for the moment the difference in extent of loss of ESR signal, offer an explanation for why phospholipid 1 (Figure Q10–3A) reacts faster with ascorbate than does phospholipid 2 (Figure Q10–3B). Note that phospholipid 1 reaches a plateau in about 15 minutes, whereas phospholipid 2 takes almost an hour.

B. To investigate the difference in extent of loss of ESR signal with the two phospholipids, the experiments were repeated using red cell ghosts that had been resealed to make them impermeable to ascorbate (Figure Q10–3C and D). Resealed red cell ghosts are missing all of their cytoplasm but have an intact plasma membrane. In these experiments the loss of ESR signal for both phospholipids was negligible in the absence of ascorbate and reached a plateau at 50% in the presence of ascorbate. What do you suppose might account for the difference in extent of loss of ESR signal in experiments with red cell ghosts (Figure Q10–3C and D) versus those with normal red cells (Figure Q10–3A and B).

C. Were the spin-labeled phospholipids introduced equally into the two monolayers of the red cell membrane?

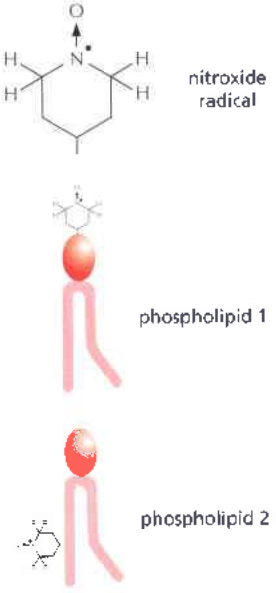


Figure Q10–2 Structures of two nitroxide-labeled lipids (Problem 10–7). The nitroxide radical is shown at the *top*, and its position of attachment to the phospholipids is shown *below*.

10–8 Monomeric single-pass transmembrane proteins span a membrane with a single α helix that has characteristic chemical properties in the region of the bilayer. Which of the three 20-amino acid sequences listed below is the most likely candidate for such a transmembrane segment? Explain the reasons for your choice. (See back of book for one-letter amino acid code; FAMILY VW is a convenient mnemonic for hydrophobic amino acids.)

A. I T L I Y F G V M A G V I G T I L L I S
B. I T P I Y F G P M A G V I G T P L L I S
C. I T E I Y F G R M A G V I G T D L L I S

10–9 You are studying the binding of proteins to the cytoplasmic face of cultured neuroblastoma cells and have found a method that gives a good yield of inside-out vesicles from the plasma membrane. Unfortunately, your preparations are contaminated with variable amounts of right-side-out vesicles. Nothing you have tried avoids this problem. A friend suggests that you pass your vesicles over an affinity column made of lectin coupled to solid beads. What is the point of your friend's suggestion?

10–10 Glycophorin, a protein in the plasma membrane of the red blood cell, normally exists as a homodimer that is held together entirely by interactions between its transmembrane domains. Since transmembrane domains are hydrophobic, how is it that they can associate with one another so specifically?

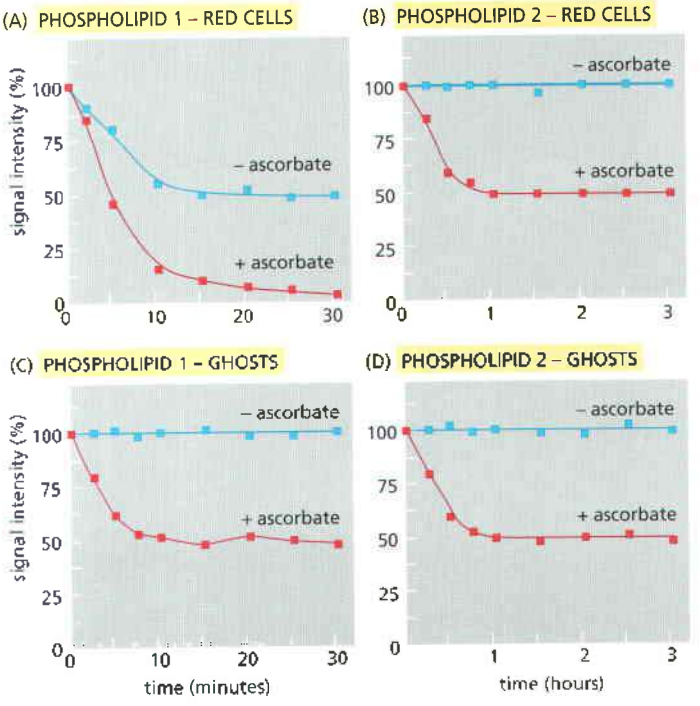


Figure Q10–3 Decrease in ESR signal intensity as a function of time in intact red cells and red cell ghosts in the presence and absence of ascorbate (Problem 10–7). (A and B) Phospholipid 1 and phospholipid 2 in intact red cells. (C and D) Phospholipid 1 and phospholipid 2 in red cell ghosts.

REFERENCES

General

Bretscher MS (1973) Membrane structure:some general principles. *Science* 181:622–629.

Edidin M (2003) Lipids on the frontier:a century of cell membrane bilayers. *Nat Rev Mol Cell Biol* 4:414–418.

Jacobson K et al (1995) Revisiting the fluid mosaic model of membranes. *Science* 268:1441–1442.

Lipowsky R & Sackmann E (eds) (1995) The structure and dynamics of membranes. Amsterdam: Elsevier.

Singer SJ & Nicolson GL (1972) The fluid mosaic model of the structure of cell membranes. *Science* 175:720–731.

The Lipid Bilayer

Bevers EM, Comfurius P & Zwaal RF (1999) Lipid translocation across the plasma membrane of mammalian cells. *Biochim Biophys Acta* 1439:317–330.

Devaux PF (1993) Lipid transmembrane asymmetry and flip-flop in biological membranes and in lipid bilayers. *Curr Opin Struct Biol* 3:489–494.

Dowhan W (1997) Molecular basis for membrane phospholipid diversity:why are there so many lipids? *Annu Rev Biochem* 66:199–232.

Hakomori S-I (2002) Inaugural Article:The glycosynapse. *Proc Natl Acad Sci USA* 99:225–232.

Harder T & Simons K (1997) Caveolae, DIGs, and the dynamics of sphingolipid-cholesterol microdomains. *Curr Opin Cell Biol* 9:534–542.

Hazel JR (1995) Thermal adaptation in biological membranes:is homeoviscous adaptation the explanation? *Annu Rev Physiol* 57:19–42.

Ichikawa S & Hirabayashi Y (1998) Glucosylceramide synthase and glycosphingolipid synthesis. *Trends Cell Biol* 8:198–202.

Kornberg RD & McConnell HM (1971) Lateral diffusion of phospholipids in a vesicle membrane. *Proc Natl Acad Sci USA* 68:2564–2568.

Mansilla MC, Cybulski LE & de Mendoza D (2004) Control of membrane lipid fluidity by molecular thermosensors. *J Bacteriol* 186:6681–6688.

McConnell HM & Radhakrishnan A (2003) Condensed complexes of cholesterol and phospholipids. *Biochim Biophys Acta* 1610:159–73.

Pomorski T & Menon AK (2006) Lipid flippases and their biological functions. *Cell Mol Life Sci* 63:2908–2921.

Rothman JE & Lenard J (1977) Membrane asymmetry. *Science* 195:743–53.

Simons K & Vaz WL (2004) Model systems, lipid rafts, and cell membranes. *Annu Rev Biophys Biomol Struct* 33:269–95.

Tanford C (1980) The Hydrophobic Effect: Formation of Micelles and Biological Membranes. New York: Wiley.

van Meer G (2005) Cellular lipidomics. *EMBO J* 24:3159–3165.

Membrane Proteins

Bennett V & Baines AJ (2001) Spectrin and ankyrin-based pathways: metazoan inventions for integrating cells into tissues. *Physiol Rev* 81:1353–1392.

Bijlmakers MJ & Marsh M (2003) The on-off story of protein palmitoylation. *Trends Cell Biol* 13:32–42.

Branden C & Tooze J (1999) Introduction to Protein Structure. 2nd ed. New York: Garland Science.

Bretscher MS & Raff MC (1975) Mammalian plasma membranes. *Nature* 258:43–49.

Buchanan SK (1999) Beta-barrel proteins from bacterial outer membranes:structure, function and refolding. *Curr Opin Struct Biol* 9:455–461.

Chen Y, Lagerholm BC & Jacobson K (2006) Methods to measure the lateral diffusion of membrane lipids and proteins. *Methods* 39:147–153.

Curranb AR & Engelman DM (2003) Sequence motifs, polar interactions and conformational changes in helical membrane proteins. *Curr Opin Struct Biol* 13:412.

Deisenhofer J & Michel H (1991) Structures of bacterial photosynthetic reaction centers. *Annu Rev Cell Biol* 7:1–23.

Drickamer K & Taylor ME (1993) Biology of animal lectins. *Annu Rev Cell Biol* 9:237–64.

Drickamer K & Taylor ME (1998) Evolving views of protein glycosylation. *Trends Biochem Sci* 23:321–324.

Frye LD & Edidin M (1970) The rapid intermixing of cell surface antigens after formation of mouse-human heterokaryons. *J Cell Sci* 7:319–335.

Helenius A and Simons K (1975) Solubilization of membranes by detergents. *Biochim Biophys Acta* 415:29–79.

Henderson R & Unwin PN (1975) Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* 257:28–32.

Kyte J & Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157:105–132.

le Maire M, Champeil P et al (2000) Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim Biophys Acta* 1508:86–111.

Lee AG (2003) Lipid-protein interactions in biological membranes:a structural perspective. *Biochim Biophys Acta* 1612:1–40.

Marchesi VT, Furthmayr H et al (1976) The red cell membrane. *Annu Rev Biochem* 45:667–698.

Nakada C, Ritchie K & Kusumi A (2003) Accumulation of anchored proteins forms membrane diffusion barriers during neuronal polarization. *Nat Cell Biol* 5:626–632.

Oesterhelt D (1998) The structure and mechanism of the family of retinal proteins from halophilic archaea. *Curr Opin Struct Biol* 8:489–500.

Rang N & van der Goot FG (2006) About lipids and toxins. *FEBS Lett* 580:5572–5579.

Reithmeier RAF (1993) The erythrocyte anion transporter (band 3) *Curr Opin Cell Biol* 7:707–714.

Rodgers W & Glaser M (1993) Distributions of proteins and lipids in the erythrocyte membrane. *Biochemistry* 32:12591–12598.

Sharon N & Lis H (2004) History of lectins:from hemagglutinins to biological recognition molecules. *Glycobiology* 14:53R–62R.

Sheetz MP (2001) Cell control by membrane-cytoskeleton adhesion. *Nat Rev Mol Cell Biol* 2:392–396.

Silvius JR (1992) Solubilization and functional reconstitution of biomembrane components. *Annu Rev Biophys Biomol Struct* 21:323–348.

Steck TL (1974) The organization of proteins in the human red blood cell membrane. A review. *J Cell Biol* 62:1–19.

Subramaniam S (1999) The structure of bacteriorhodopsin:an emerging consensus. *Curr Opin Struct Biol* 9:462–468.

Viel A & Branton D (1996) Spectrin:on the path from structure to function. *Curr Opin Cell Biol* 8:49–55.

Wallin E & von Heijne G (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci* 7:1029–1038.

White SH & Wimley WC (1999) Membrane protein folding and stability:physical principles. *Annu Rev Biophys Biomol Struct* 28:319–365.

Membrane Transport of Small Molecules and the Electrical Properties of Membranes

11

Because of its hydrophobic interior, the lipid bilayer of cell membranes prevents the passage of most polar molecules. This barrier function allows the cell to maintain concentrations of solutes in its cytosol that differ from those in the extracellular fluid and in each of the intracellular membrane-enclosed compartments. To benefit from this barrier, however, cells have had to evolve ways of transferring specific water-soluble molecules and ions across their membranes in order to ingest essential nutrients, excrete metabolic waste products, and regulate intracellular ion concentrations. Cells use specialized transmembrane proteins to transport inorganic ions and small water-soluble organic molecules across the lipid bilayer. Cells can also transfer macromolecules and even large particles across their membranes, but the mechanisms involved in most of these cases differ from those used for transferring small molecules, and they are discussed in Chapters 12 and 13. The importance of membrane transport is reflected in the large number of genes in all organisms that code for transport proteins, which make up 15–30% of the membrane proteins in all cells. Some specialized mammalian cells devote up to two-thirds of their total metabolic energy consumption to membrane transport processes.

We begin this chapter by describing some general principles of how small water-soluble molecules traverse cell membranes. We then consider, in turn, the two main classes of membrane proteins that mediate this traffic of molecules back and forth across lipid bilayers: *transporters*, which have moving parts to transport specific molecules across membranes, and *channels*, which form a narrow hydrophilic pore, allowing passive transmembrane movement, primarily of small inorganic ions. Transporters can be coupled to a source of energy to catalyze active transport, and a combination of selective passive permeability and active transport creates large differences in the composition of the cytosol compared with that of either the extracellular fluid (**Table 11–1**) or the fluid within membrane-enclosed organelles. By generating ionic concentration differences across the lipid bilayer, cell membranes can store potential energy in the form of electrochemical gradients, which drive various transport processes, convey electrical signals in electrically excitable cells, and (in mitochondria, chloroplasts, and bacteria) make most of the cell's ATP. We focus our discussion mainly on transport across the plasma membrane, but similar mechanisms operate across the other membranes of the eucaryotic cell, as discussed in later chapters.

In the last part of the chapter, we concentrate mainly on the functions of ion channels in neurons (nerve cells). In these cells, channel proteins perform at their highest level of sophistication, enabling networks of neurons to carry out all the human brain's astonishing feats.

In This Chapter

PRINCIPLES OF MEMBRANE TRANSPORT 651

TRANSPORTERS AND ACTIVE MEMBRANE TRANSPORT 654

ION CHANNELS AND THE ELECTRICAL PROPERTIES OF MEMBRANES 667

PRINCIPLES OF MEMBRANE TRANSPORT

We begin this section by describing the permeability properties of protein-free, synthetic lipid bilayers. We then introduce some of the terms used to describe the various forms of membrane transport and some strategies for characterizing the proteins and processes involved.

Table 11–1 A Comparison of Ion Concentrations Inside and Outside a Typical Mammalian Cell

COMPONENT	INTRACELLULAR CONCENTRATION (mM)	EXTRACELLULAR CONCENTRATION (mM)
Cations		
Na^+	5–15	145
K^+	140	5
Mg^{2-}	0.5	1–2
Ca^{2+}	10^{-4}	1–2
H^+	7×10^{-5} ($10^{-7.2}$ M or pH 7.2)	4×10^{-5} ($10^{-7.4}$ M or pH 7.4)
Anions*		
Cl^-	5–15	110

*The cell must contain equal quantities of positive and negative charges (that is, it must be electrically neutral). Thus, in addition to Cl^-, the cell contains many other anions not listed in this table; in fact, most cell constituents are negatively charged (HCO_3^-, PO_4^{3-}, proteins, nucleic acids, metabolites carrying phosphate and carboxyl groups, etc.). The concentrations of Ca^{2+} and Mg^{2+} given are for the free ions. There is a total of about 20 mM Mg^{2+} and 1–2 mM Ca^{2+} in cells, but both are mostly bound to proteins and other substances and, for Ca^{2+}, stored within various organelles.

Protein-Free Lipid Bilayers Are Highly Impermeable to Ions

Given enough time, virtually any molecule will diffuse across a protein-free lipid bilayer down its concentration gradient. The rate of diffusion, however, varies enormously, depending partly on the size of the molecule but mostly on its relative solubility in oil. In general, the smaller the molecule and the more soluble it is in oil (the more hydrophobic, or nonpolar, it is), the more rapidly it will diffuse across a lipid bilayer. Small nonpolar molecules, such as O_2 and CO_2, readily dissolve in lipid bilayers and therefore diffuse rapidly across them. Small uncharged polar molecules, such as water or urea, also diffuse across a bilayer, albeit much more slowly (**Figure 11–1**). By contrast, lipid bilayers are highly impermeable to charged molecules (ions), no matter how small: the charge and high degree of hydration of such molecules prevents them from entering the hydrocarbon phase of the bilayer. Thus, synthetic lipid bilayers are 10^9 times more permeable to water than to even such small ions as Na^+ or K^+ (**Figure 11–2**).

There Are Two Main Classes of Membrane Transport Proteins: Transporters and Channels

Like synthetic lipid bilayers, cell membranes allow water and nonpolar molecules to permeate by simple diffusion. Cell membranes, however, also have to allow the passage of various polar molecules, such as ions, sugars, amino acids, nucleotides, and many cell metabolites that cross synthetic lipid bilayers only very slowly. Special **membrane transport proteins** transfer such solutes across cell membranes. These proteins occur in many forms and in all types of biological membranes. Each protein transports a particular class of molecule (such as ions, sugars, or amino acids) and often only certain molecular species of the class. Studies in the 1950s found that bacteria with a single-gene mutation were unable to transport sugars across their plasma membrane, thereby demonstrating the specificity of membrane transport proteins. We now know that humans with similar mutations suffer from various inherited diseases that hinder the transport of a specific solute in the kidney, intestine, or other cell type. Individuals with the inherited disease *cystinuria*, for example, cannot transport certain amino acids (including cystine, the disulfide-linked dimer of cysteine) from either the urine or the intestine into the blood; the resulting accumulation of cystine in the urine leads to the formation of cystine stones in the kidneys.

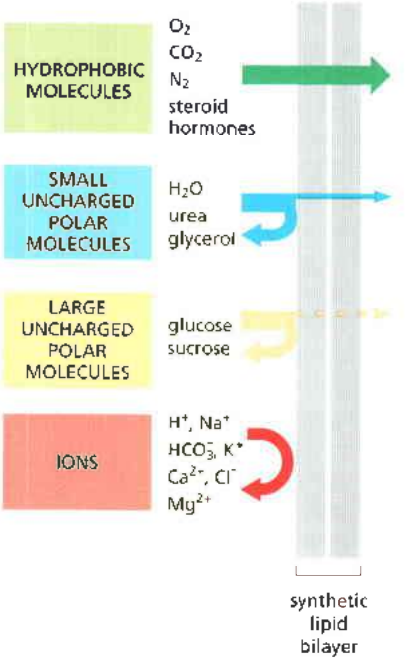


Figure 11–1 The relative permeability of a synthetic lipid bilayer to different classes of molecules. The smaller the molecule and, more importantly, the less strongly it associates with water, the more rapidly the molecule diffuses across the bilayer.

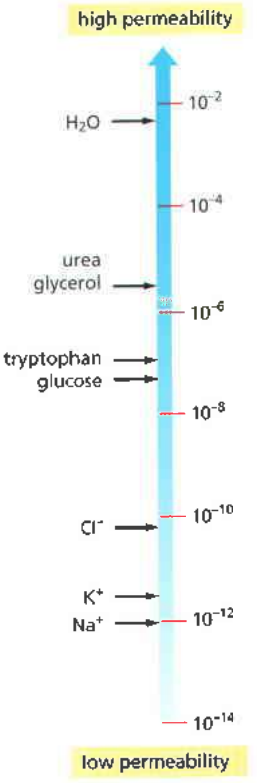


Figure 11–2 Permeability coefficients for the passage of various molecules through synthetic lipid bilayers. The rate of flow of a solute across the bilayer is directly proportional to the difference in its concentration on the two sides of the membrane. Multiplying this concentration difference (in mol/cm^3) by the permeability coefficient (in cm/sec) gives the flow of solute in moles per second per square centimeter of bilayer. A concentration difference of tryptophan of 10^{-4} mol/cm^3 ($10^{-4}/10^{-3}$ L = 0.1 M), for example, would cause a flow of 10^{-4} mol/cm^3 $\times$ 10^{-7} cm/sec = 10^{-11} mol/sec through 1 cm^2 of bilayer, or 6×10^4 molecules/sec through 1 μm^2 of bilayer.

All membrane transport proteins that have been studied in detail have been found to be multipass transmembrane proteins—that is, their polypeptide chains traverse the lipid bilayer multiple times. By forming a continuous protein pathway across the membrane, these proteins enable specific hydrophilic solutes to cross the membrane without coming into direct contact with the hydrophobic interior of the lipid bilayer.

Transporters and channels are the two major classes of membrane transport proteins (**Figure 11–3**). **Transporters** (also called *carriers*, or *permeases*) bind the specific solute to be transported and undergo a series of conformational changes to transfer the bound solute across the membrane. **Channels**, in contrast, interact with the solute to be transported much more weakly. They form aqueous pores that extend across the lipid bilayer; when open, these pores allow specific solutes (usually inorganic ions of appropriate size and charge) to pass through them and thereby cross the membrane. Not surprisingly, transport through channels occurs at a much faster rate than transport mediated by transporters. Although water can diffuse across synthetic lipid bilayers, all cells contain specific channel proteins (called *water channels*, or *aquaporins*) that greatly increase the permeability of these membranes to water, as we discuss later.

Active Transport Is Mediated by Transporters Coupled to an Energy Source

All channels and many transporters allow solutes to cross the membrane only passively ("downhill"), a process called **passive transport**, or **facilitated diffusion**. In the case of transport of a single uncharged molecule, the difference in the concentration on the two sides of the membrane—its *concentration gradient*—drives passive transport and determines its direction (**Figure 11–4A**).

If the solute carries a net charge, however, both its concentration gradient and the electrical potential difference across the membrane, the *membrane potential*, influence its transport. The concentration gradient and the electrical gradient combine to form a net driving force, the **electrochemical gradient**, for each charged solute (Figure 11–4B). We discuss electrochemical gradients in more detail in Chapter 14. In fact, almost all plasma membranes have an electrical potential difference (voltage gradient) across them, with the inside usually negative with respect to the outside. This potential difference favors the entry of positively charged ions into the cell but opposes the entry of negatively charged ions.

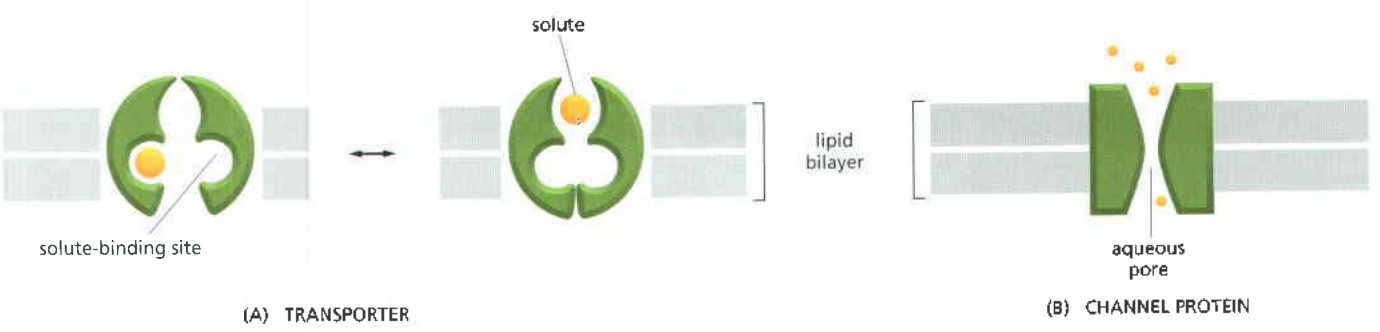


Figure 11–3 Transporters and channel proteins. (A) A transporter alternates between two conformations, so that the solute-binding site is sequentially accessible on one side of the bilayer and then on the other. (B) In contrast, a channel protein forms a water-filled pore across the bilayer through which specific solutes can diffuse.

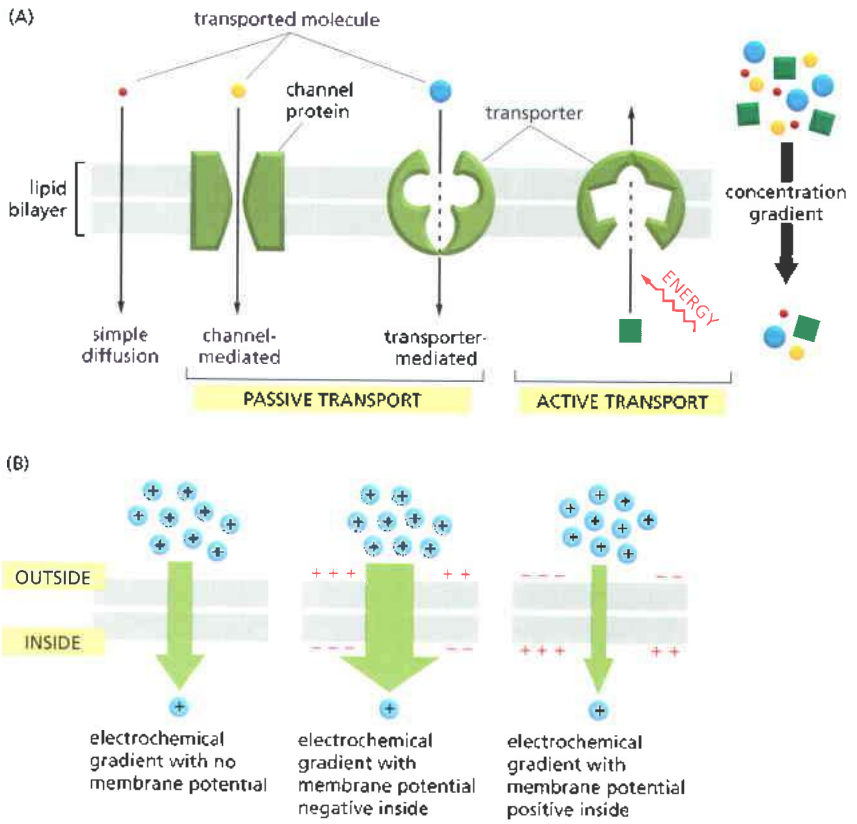


Figure 11–4 Passive and active transport compared. (A) Passive transport down an electrochemical gradient occurs spontaneously, either by simple diffusion through the lipid bilayer or by facilitated diffusion through channels and passive transporters. By contrast, active transport requires an input of metabolic energy and is always mediated by transporters that harvest metabolic energy to pump the solute against its electrochemical gradient. (B) An electrochemical gradient combines the membrane potential and the concentration gradient; they can work additively to increase the driving force on an ion across the membrane *(middle)* or can work against each other *(right)*.

Cells also require transport proteins that will actively pump certain solutes across the membrane against their electrochemical gradients ("uphill"); this process, known as **active transport**, is mediated by transporters, which are also called *pumps*. In active transport, the pumping activity of the transporter is directional because it is tightly coupled to a source of metabolic energy, such as ATP hydrolysis or an ion gradient, as discussed later. Thus, transmembrane movement of small molecules mediated by transporters can be either active or passive, whereas that mediated by channels is always passive.

Summary

Lipid bilayers are highly impermeable to most polar molecules. To transport small water-soluble molecules into or out of cells or intracellular membrane-enclosed compartments, cell membranes contain various membrane transport proteins, each of which is responsible for transferring a particular solute or class of solutes across the membrane. There are two classes of membrane transport proteins—transporters and channels. Both form continuous protein pathways across the lipid bilayer. Whereas transmembrane movement mediated by transporters can be either active or passive, solute flow through channel proteins is always passive.

TRANSPORTERS AND ACTIVE MEMBRANE TRANSPORT

The process by which a transporter transfers a solute molecule across the lipid bilayer resembles an enzyme–substrate reaction, and in many ways transporters behave like enzymes. In contrast to ordinary enzyme–substrate reactions, however, the transporter does not modify the transported solute but instead delivers it unchanged to the other side of the membrane.

Each type of transporter has one or more specific binding sites for its solute (substrate). It transfers the solute across the lipid bilayer by undergoing

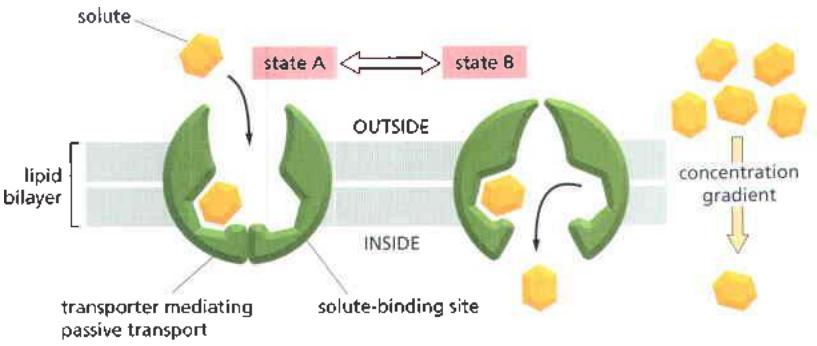


Figure 11–5 A model of how a conformational change in a transporter could mediate the passive movement of a solute. The transporter shown can exist in two conformational states: in state A, the binding sites for solute are exposed on the outside of the lipid bilayer; in state B, the same sites are exposed on the other side of the bilayer. The transition between the two states can occur randomly. It is completely reversible and does not depend on whether the solute binding site is occupied. Therefore, if the solute concentration is higher on the outside of the bilayer, more solute binds to the transporter in the A conformation than in the B conformation, and there is a net transport of solute down its concentration gradient (or, if the solute is an ion, down its electrochemical gradient).

reversible conformational changes that alternately expose the solute-binding site first on one side of the membrane and then on the other. Figure 11–5 shows a schematic model of how a transporter operates. When the transporter is saturated (that is, when all solute-binding sites are occupied), the rate of transport is maximal. This rate, referred to as V_{max} (V for velocity), is characteristic of the specific carrier. V_{max} measures the rate with which the carrier can flip between its two conformational states. In addition, each transporter has a characteristic affinity for its solute, reflected in the K_m of the reaction, which is equal to the concentration of solute when the transport rate is half its maximum value (Figure 11–6). As with enzymes, the binding of solute can be blocked specifically by either competitive inhibitors (which compete for the same binding site and may or may not be transported) or noncompetitive inhibitors (which bind elsewhere and specifically alter the structure of the transporter).

As we discuss below, it requires only a relatively minor modification of the model shown in Figure 11–5 to link a transporter to a source of energy in order to pump a solute uphill against its electrochemical gradient. Cells carry out such active transport in three main ways (Figure 11–7):

1. *Coupled transporters* couple the uphill transport of one solute across the membrane to the downhill transport of another.
2. *ATP-driven pumps* couple uphill transport to the hydrolysis of ATP.
3. *Light-driven pumps*, which are found mainly in bacteria and archaea, couple uphill transport to an input of energy from light, as with bacteriorhodopsin (discussed in Chapter 10).

Amino acid sequence comparisons suggest that, in many cases, there are strong similarities in molecular design between transporters that mediate active transport and those that mediate passive transport. Some bacterial transporters, for example that use the energy stored in the H^+ gradient across the plasma membrane to drive the active uptake of various sugars are structurally similar to the transporters that mediate passive glucose transport into most animal cells. This suggests an evolutionary relationship between various transporters. Given the importance of small metabolites and sugars as energy sources, it is not surprising that the superfamily of transporters is an ancient one.

We begin our discussion of active transport by considering transporters that are driven by ion gradients. These proteins have a crucial role in the transport of small metabolites across membranes in all cells. We then discuss ATP-driven pumps, including the Na^+ pump that is found in the plasma membrane of almost all cells.

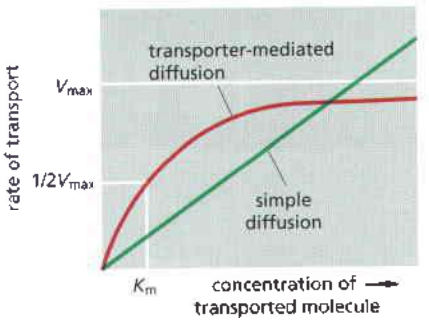


Figure 11–6 The kinetics of simple diffusion and transporter-mediated diffusion. Whereas the rate of simple diffusion is always proportional to the solute concentration, the rate of transporter-mediated diffusion reaches a maximum (V_{max}) when the transporter is saturated. The solute concentration when transport is at half its maximal value approximates the binding constant (K_m) of the transporter for the solute and is analogous to the K_m of an enzyme for its substrate. The graph applies to a transporter moving a single solute; the kinetics of coupled transport of two or more solutes is more complex.

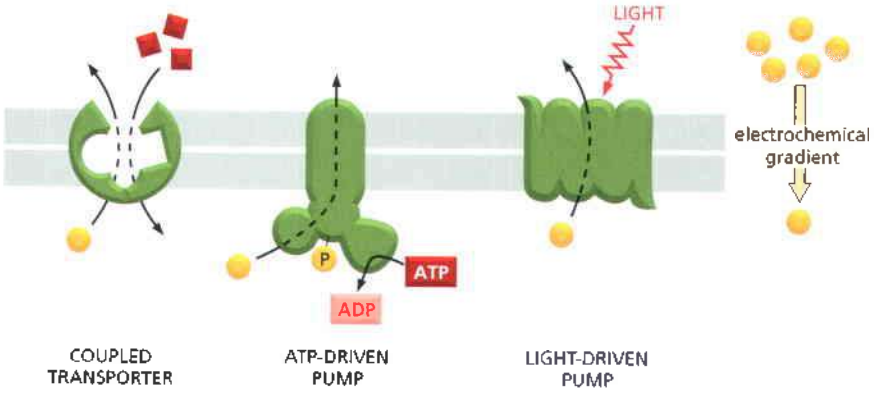


Figure 11–7 **Three ways of driving active transport.** The actively transported molecule is shown in *yellow*, and the energy source is shown in *red*.

Active Transport Can Be Driven by Ion Gradients

Some transporters simply mediate the movement of a single solute from one side of the membrane to the other at a rate determined by their V_{max} and K_m; they are called **uniporters**. Others function as *coupled transporters*, in which the transfer of one solute strictly depends on the transport of a second. Coupled transport involves either the simultaneous transfer of a second solute in the same direction, performed by **symporters** (also called *co-transporters*), or the transfer of a second solute in the opposite direction, performed by **antiporters** (also called *exchangers*) (**Figure 11–8**).

The tight coupling between the transfer of two solutes allows these coupled transporters to harvest the energy stored in the electrochemical gradient of one solute, typically an ion, to transport the other. In this way, the free energy released during the movement of an inorganic ion down an electrochemical gradient is used as the driving force to pump other solutes uphill, against their electrochemical gradient. This principle can work in either direction; some coupled transporters function as symporters, others as antiporters. In the plasma membrane of animal cells, Na^+ is the usual co-transported ion, the electrochemical gradient of which provides a large driving force for the active transport of a second molecule. The Na^+ that enters the cell during transport is subsequently pumped out by an ATP-driven Na^+ pump in the plasma membrane (as we discuss later), which, by maintaining the Na^+ gradient, indirectly drives the transport. (For this reason ion-driven carriers are said to mediate *secondary active transport*, whereas ATP-driven carriers are said to mediate *primary active transport*.)

Intestinal and kidney epithelial cells, for example, contain a variety of symporters that are driven by the Na^+ gradient across the plasma membrane. Each Na^+-driven symporter is specific for importing a small group of related sugars or amino acids into the cell, and the solute and Na^+ bind to different sites on the transporter. Because the Na^+ tends to move into the cell down its electrochemical gradient, the sugar or amino acid is, in a sense, "dragged" into the cell with it. The greater the electrochemical gradient for Na^+, the greater the rate of solute

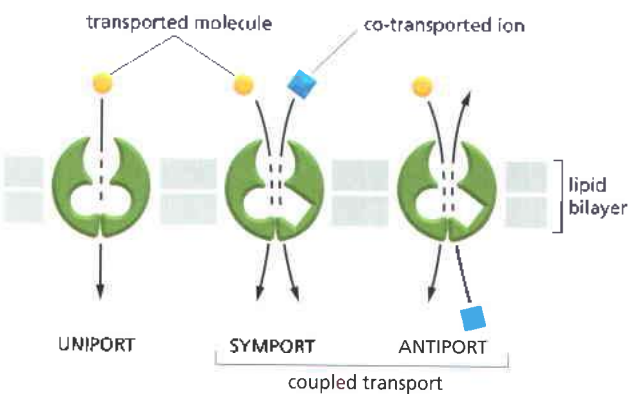


Figure 11–8 **Three types of transporter-mediated movement.** <ACCC> This schematic diagram shows transporters functioning as uniporters, symporters, and antiporters.

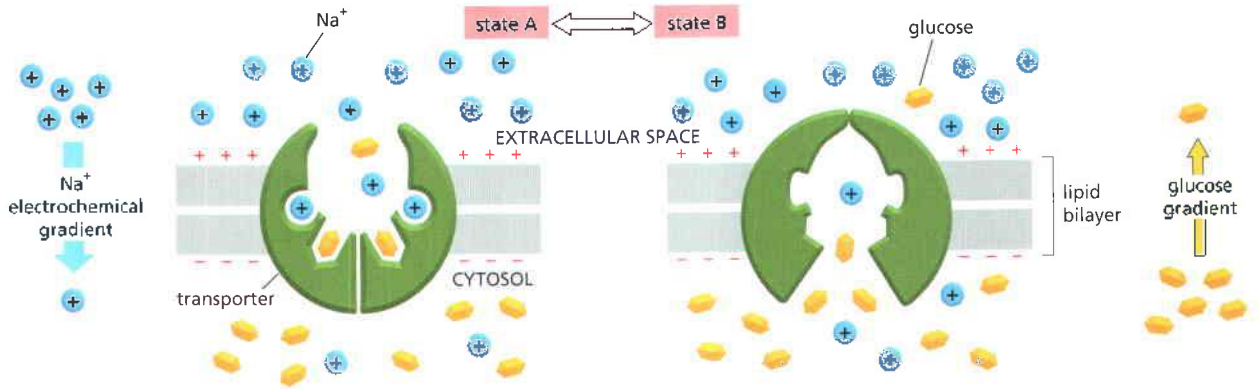


Figure 11–9 One way in which a glucose transporter can be driven by a Na^+ gradient. As in the model shown in Figure 11–5, the transporter oscillates between two alternate states, A and B. In the A state, the protein is open to the extracellular space; in the B state, it is open to the cytosol. Binding of Na^+ and glucose is cooperative—that is, the binding of either ligand induces a conformational change that increases the protein's affinity for the other ligand. Since the Na^+ concentration is much higher in the extracellular space than in the cytosol, glucose is more likely to bind to the transporter in the A state. Therefore, both Na^+ and glucose enter the cell (via an A → B transition) much more often than they leave it (via a B → A transition). The overall result is the net transport of both Na^+ and glucose into the cell. Note that, because the binding is cooperative, if one of the two solutes is missing, the other fails to bind to the transporter. Thus, the transporter undergoes a conformational switch between the two states only if both solutes or neither are bound.

entry; conversely, if the Na^+ concentration in the extracellular fluid is reduced, solute transport decreases (**Figure 11–9**).

In bacteria and yeasts, as well as in many membrane-enclosed organelles of animal cells, most active transport systems driven by ion gradients depend on H^+ rather than Na^+ gradients, reflecting the predominance of H^+ pumps and the virtual absence of Na^+ pumps in these membranes. The electrochemical H^+ gradient drives the active transport of many sugars and amino acids across the plasma membrane and into bacterial cells. One well-studied H^+-driven symporter is **lactose permease**, which transports lactose across the plasma membrane of *E. coli*. Structural and biophysical studies of the permease, as well as extensive analyses of mutant forms of the protein, have led to a detailed model of how the symporter works. The permease consists of 12 loosely packed transmembrane α helices. During the transport cycle, some of the helices undergo sliding motions that cause them to tilt. These motions alternately open and close a crevice between the helices, exposing the binding sites for lactose and H^+, first on one side of the membrane and then on the other (**Figure 11–10**).

Transporters in the Plasma Membrane Regulate Cytosolic pH

Most proteins operate optimally at a particular pH. Lysosomal enzymes, for example, function best at the low pH (~5) found in lysosomes, whereas cytosolic enzymes function best at the close to neutral pH (~7.2) found in the cytosol. It is therefore crucial that cells control the pH of their intracellular compartments.

Most cells have one or more types of Na^+-driven antiporters in their plasma membrane that help to maintain the cytosolic pH at about 7.2. These transporters use the energy stored in the Na^+ gradient to pump out excess H^+, which either leaks in or is produced in the cell by acid-forming reactions. Two mechanisms are used: either H^+ is directly transported out of the cell or HCO_3^- is brought into the cell to neutralize H^+ in the cytosol (according to the reaction $HCO_3^- + H^+ \rightarrow H_2O + CO_2$). One of the antiporters that uses the first mechanism is a *Na^+–H^+ exchanger*, which couples an influx of Na^+ to an efflux of H^+. Another, which uses a combination of the two mechanisms, is a *Na^+-driven Cl^-–HCO_3^- exchanger* that couples an influx of Na^+ and HCO_3^- to an efflux of Cl^- and H^+ (so

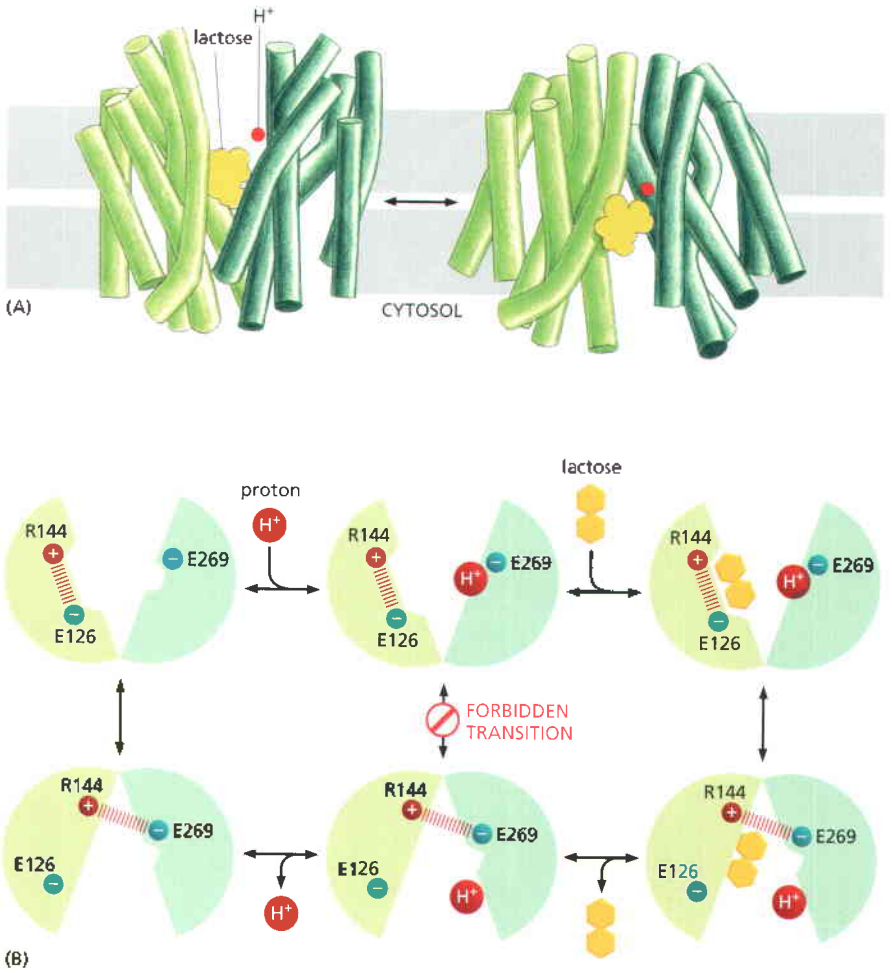


Figure 11–10 The molecular mechanism of the bacterial lactose permease suggested from its crystal structure. (A) The 12 transmembrane helices of the permease are clustered into two lobes, shown in two shades of *green*. The loops that connect the helices on either side of the membrane are omitted for clarity. During transport, the helices slide and tilt in the membrane, exposing binding sites for the disaccharide lactose *(yellow)* and H^+ to either side of the membrane. (B) In one conformational state, the H^+- and lactose-binding sites are accessible to the extracellular space *(top row)*; in the other, they are exposed to the cytosol *(bottom row)*. Loading the solutes on the extracellular side is favored because arginine (R) 144 forms a bond with glutamic acid (E) 126, leaving E269 free to accept H^+. Unloading the solutes on the cytosolic side is favored because R144 forms a bond with E269, which destabilizes the bound H^+. In addition, the lactose-binding site is partially disrupted due to the rearrangement of the helices. Because the transition between the two protonated states *(middle)* is forbidden, H^+ can only be transported when a lactose is also transported. In this way, the electrochemical H^+ gradient drives lactose import. (Adapted from J. Abramson et al., *Science* 301: 610–615, 2003. With permission from AAAS.)

that $NaHCO_3$ comes in and HCl goes out). The Na^+-driven Cl^-–HCO_3^- exchanger is twice as effective as the Na^+–H^+ exchanger: it pumps out one H^+ and neutralizes another for each Na^+ that enters the cell. If HCO_3^- is available, as is usually the case, this antiporter is the most important transporter regulating the cytosolic pH. The pH inside the cell regulates both exchangers; when the pH in the cytosol falls, both exchangers increase their activity.

A *Na^+-independent Cl^-–HCO_3^- exchanger* adjusts the cytosolic pH in the reverse direction. Like the Na^+-dependent transporters, pH regulates the Na^+-independent Cl^-–HCO_3^- exchanger, but the exchanger's activity increases as the cytosol becomes too alkaline. The movement of HCO_3^- in this case is normally out of the cell, down its electrochemical gradient, which decreases the pH of the cytosol. A Na^+-independent Cl^-–HCO_3^- exchanger in the membrane of red blood cells (called band 3 protein—see Figure 10–41) facilitates the quick discharge of CO_2 (as HCO_3^-) as the cells pass through capillaries in the lung.

The intracellular pH is not entirely regulated by these coupled transporters: ATP-driven H^+ pumps are also used to control the pH of many intracellular compartments. As discussed in Chapter 13, H^+ pumps maintain the low pH in lysosomes, as well as in endosomes and secretory vesicles. These H^+ pumps use the energy of ATP hydrolysis to pump H^+ into these organelles from the cytosol.

An Asymmetric Distribution of Transporters in Epithelial Cells Underlies the Transcellular Transport of Solutes

In epithelial cells, such as those that absorb nutrients from the gut, transporters are distributed nonuniformly in the plasma membrane and thereby contribute to the **transcellular transport** of absorbed solutes. By the action of

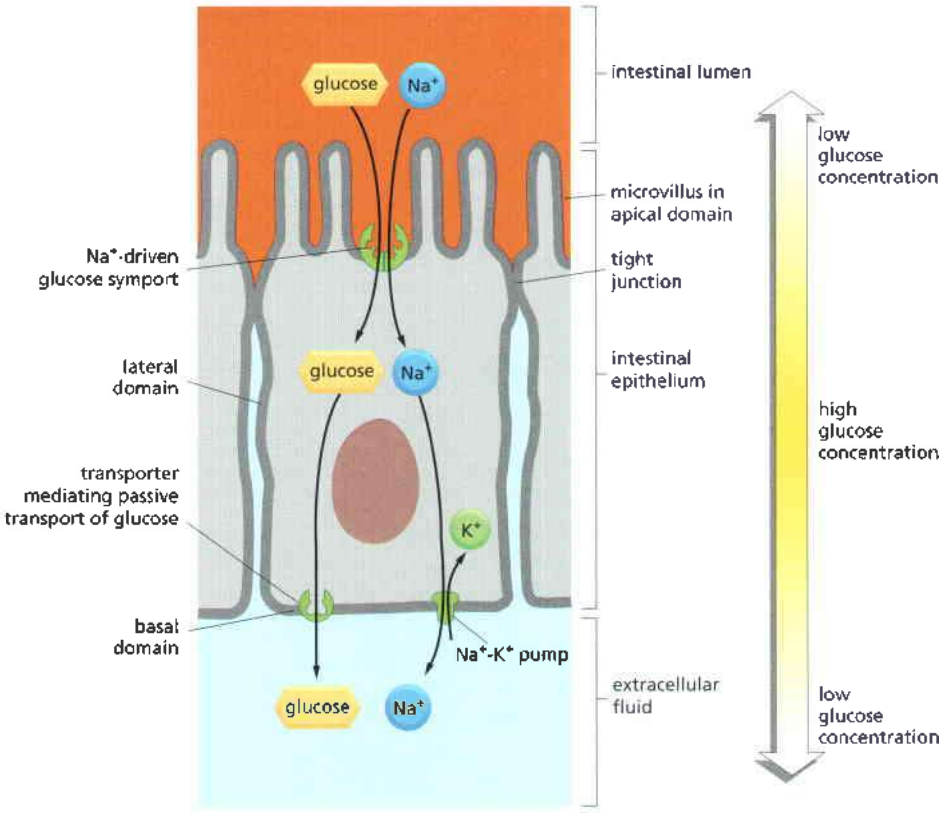


Figure 11–11 Transcellular transport. <GGAT> The transcellular transport of glucose across an intestinal epithelial cell depends on the nonuniform distribution of transporters in the cell's plasma membrane. The process shown here results in the transport of glucose from the intestinal lumen to the extracellular fluid (from where it passes into the blood). Glucose is pumped into the cell through the apical domain of the membrane by a Na^+-powered glucose symporter. Glucose passes out of the cell (down its concentration gradient) by passive movement through a different glucose transporter in the basal and lateral membrane domains. The Na^+ gradient driving the glucose symport is maintained by a Na^+ pump in the basal and lateral plasma membrane domains, which keeps the internal concentration of Na^+ low. Adjacent cells are connected by impermeable tight junctions, which have a dual function in the transport process illustrated: they prevent solutes from crossing the epithelium between cells, allowing a concentration gradient of glucose to be maintained across the cell sheet, and they also serve as diffusion barriers within the plasma membrane, which help confine the various transporters to their respective membrane domains (see Figure 10–37).

the transporters in these cells, solutes are moved across the epithelial cell layer into the extracellular fluid from where they pass into the blood. As shown in **Figure 11–11**, Na^+-linked symporters located in the apical (absorptive) domain of the plasma membrane actively transport nutrients into the cell, building up substantial concentration gradients for these solutes across the plasma membrane. Na^+-independent transport proteins in the basal and lateral (basolateral) domain allow the nutrients to leave the cell passively down these concentration gradients.

In many of these epithelial cells, the plasma membrane area is greatly increased by the formation of thousands of microvilli, which extend as thin, fingerlike projections from the apical surface of each cell. Such microvilli can increase the total absorptive area of a cell as much as 25-fold, thereby enhancing its transport capabilities.

As we have seen, ion gradients have a crucial role in driving many essential transport processes in cells. Ion pumps that use the energy of ATP hydrolysis establish and maintain these gradients, as we discuss next.

There Are Three Classes of ATP-Driven Pumps

ATP-driven pumps are often called *transport ATPases* because they hydrolyze ATP to ADP and phosphate and use the energy released to pump ions or other solutes across a membrane. There are three principal classes of ATP-driven pumps (**Figure 11–12**), and representatives of each are found in all procaryotic and eucaryotic cells.

1. **P-type pumps** are structurally and functionally related multipass transmembrane proteins. They are called "P-type" because they phosphorylate themselves during the pumping cycle. This class includes many of the ion pumps that are responsible for setting up and maintaining gradients of Na^+, K^+, H^+, and Ca^{2+} across cell membranes.
2. **F-type pumps** are turbine-like proteins, constructed from multiple different subunits. They differ structurally from P-type ATPases and are found in the plasma membrane of bacteria, the inner membrane of mitochondria,

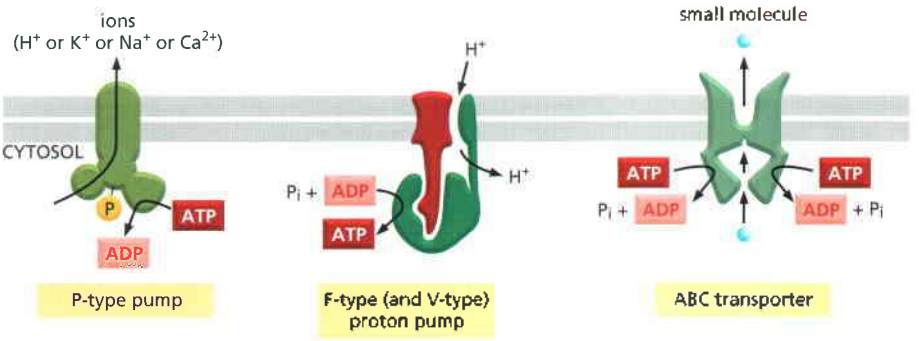


Figure 11–12 **Three types of ATP-driven pumps.** The different molecular designs of the pumps are cartooned here. Like any enzyme, pumps can work in reverse: when the electrochemical gradients of the solutes are reversed and the ATP/ADP ratio is low, they can synthesize ATP from ADP, as shown for the F-type ATPase, which normally works in this mode.

and the thylakoid membrane of chloroplasts. They are often called *ATP synthases* because they normally work in reverse: instead of using ATP hydrolysis to drive H^+ transport, they use the H^+ gradient across the membrane to drive the synthesis of ATP from ADP and phosphate. The H^+ gradient is generated either during the electron-transport steps of oxidative phosphorylation (in aerobic bacteria and mitochondria), during photosynthesis (in chloroplasts), or by the light-activated H^+ pump (bacteriorhodopsin) in *Halobacterium*. We discuss these proteins in detail in Chapter 14.

Structurally related to the F-type ATPases is a distinct family of *V-type ATPases* that normally pump H^+ rather than synthesize ATP. They pump H^+ into organelles, such as lysosomes, synaptic vesicles, and plant vacuoles to acidify the interior of these organelles (see Figure 13–36).

3. **ABC transporters** primarily pump small molecules across cell membranes, in contrast to P-type and the F- or V-type ATPases, which exclusively transport ions.

For the remainder of this section, we focus on P-type pumps and ABC transporters.

The Ca^{2+} Pump Is the Best-Understood P-type ATPase

Eucaryotic cells maintain very low concentrations of free Ca^{2+} in their cytosol ($\sim 10^{-7}$ M) in the face of a very much higher extracellular Ca^{2+} concentration ($\sim 10^{-3}$ M). Even a small influx of Ca^{2+} significantly increases the concentration of free Ca^{2+} in the cytosol, and the flow of Ca^{2+} down its steep concentration gradient in response to extracellular signals is one means of transmitting these signals rapidly across the plasma membrane (discussed in Chapter 15). It is important, therefore, that the cell maintain a steep Ca^{2+} gradient across its plasma membrane. Ca^{2+} transporters that actively pump Ca^{2+} out of the cell help maintain the gradient. One of these is a P-type Ca^{2+} ATPase; the other is an antiporter (called a *Na^+–Ca^{2+} exchanger*) that is driven by the Na^+ electrochemical gradient across the membrane (see Figure 15–41).

The best-understood P-type transport ATPase is the **Ca^{2+} pump**, or **Ca^{2+} ATPase**, in the *sarcoplasmic reticulum* (SR) membrane of skeletal muscle cells. The SR is a specialized type of endoplasmic reticulum that forms a network of tubular sacs in the muscle cell cytoplasm and serves as an intracellular store of Ca^{2+}. (When an action potential depolarizes the muscle cell plasma membrane, Ca^{2+} is released into the cytosol from the SR through *Ca^{2+}-release channels*, stimulating the muscle to contract, as discussed in Chapter 16.) The Ca^{2+} pump, which accounts for about 90% of the membrane protein of the SR, moves Ca^{2+} from the cytosol back into the SR. The endoplasmic reticulum of nonmuscle cells contains a similar Ca^{2+} pump, but in smaller quantities.

The three-dimensional structure of the SR Ca^{2+} pump has been determined by x-ray crystallography. This structure and the analysis of a related fungal H^+ pump have provided the first views of P-type transport ATPases, which are all thought to have similar structures. They contain 10 transmembrane α helices, three of which line a central channel that spans the lipid bilayer. In the unphosphorylated state, two helices are disrupted and form a cavity that binds two Ca^{2+}

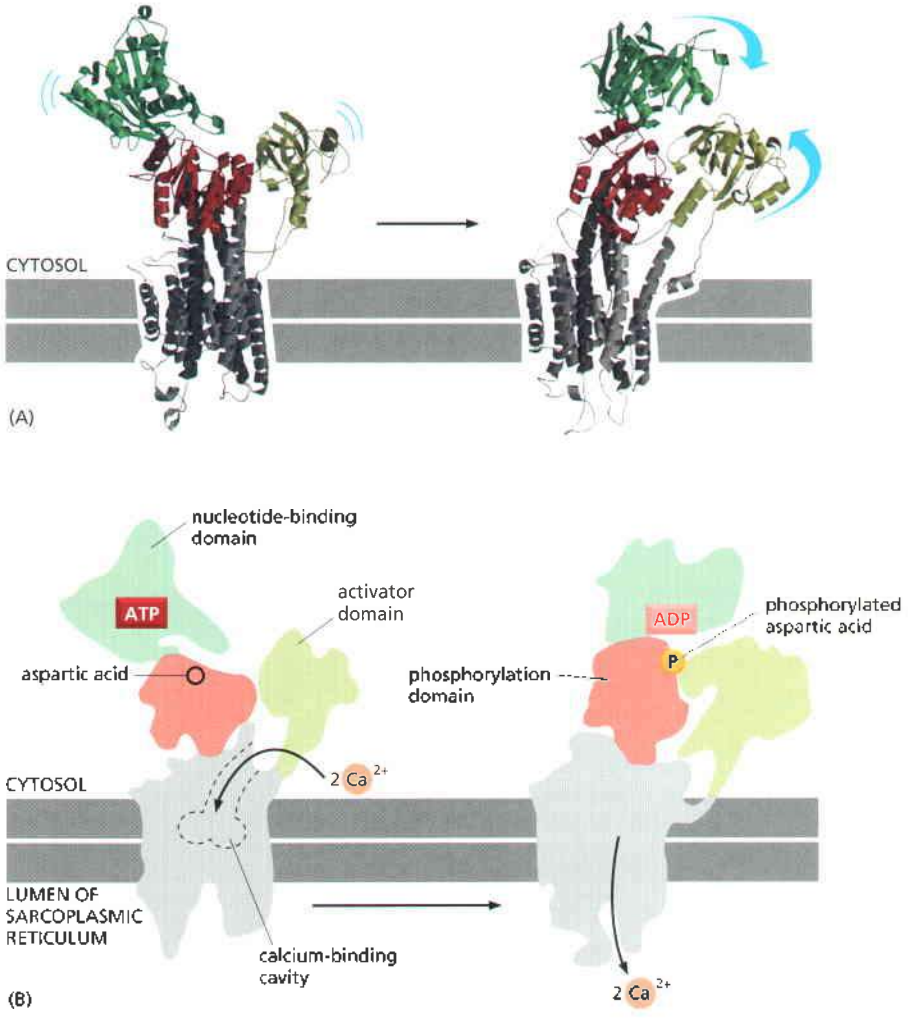


Figure 11–13 A model of how the sarcoplasmic reticulum Ca^{2+} pump moves Ca^{2+}. (A) The structures of the unphosphorylated, Ca^{2+}-bound state *(left)* and the phosphorylated, Ca^{2+}-free state *(right)* were determined by x-ray crystallography. (B) The model shows how ATP binding and hydrolysis cause drastic conformational changes, bringing the nucleotide-binding and phosphorylation domains into close proximity. This change is thought to cause a 90° rotation of the activator domain, which leads to a rearrangement of the transmembrane helices. The rearrangement of the helices disrupts the Ca^{2+}-binding cavity and releases the Ca^{2+} into the lumen of the sarcoplasmic reticulum. (Adapted from C. Toyoshima et al., *Nature* 405:647–655, 2000. With permission from Macmillan Publishers Ltd.)

ions and is accessible from the cytosolic side of the membrane. The binding of ATP to a binding site on the same side of the membrane and the subsequent transfer of the terminal phosphate group of the ATP to an aspartic acid of an adjacent domain lead to a drastic rearrangement of the transmembrane helices. The rearrangement disrupts the Ca^{2+}-binding site and releases the Ca^{2+} ions on the other side of the membrane, into the lumen of the SR (**Figure 11–13**). An essential characteristic of all P-type pumps is that the pump transiently phosphorylates itself during the pumping cycle.

The Plasma Membrane P-type Na^+-K^+ Pump Establishes the Na^+ Gradient Across the Plasma Membrane

The concentration of K^+ is typically 10–30 times higher inside cells than outside, whereas the reverse is true of Na^+ (see Table 11–1, p. 652). A **Na^+-K^+ pump**, or **Na^+ pump**, found in the plasma membrane of virtually all animal cells, maintains these concentration differences. The pump operates as an ATP-driven antiporter, actively pumping Na^+ out of the cell against its steep electrochemical gradient and pumping K^+ in (**Figure 11–14**). Because the pump hydrolyzes ATP to pump Na^+ out and K^+ in, it is also known as a **Na^+-K^+ ATPase**. The pump belongs to the family of P-type ATPases and functions very similarly to the Ca^{2+} pump (**Figure 11–15**).

We mentioned earlier that the Na^+ gradient produced by the Na^+-K^+ pump drives the transport of most nutrients into animal cells and also has a crucial role in regulating cytosolic pH. A typical animal cell devotes almost one-third of its energy to fueling this pump, and the pump consumes even more energy in

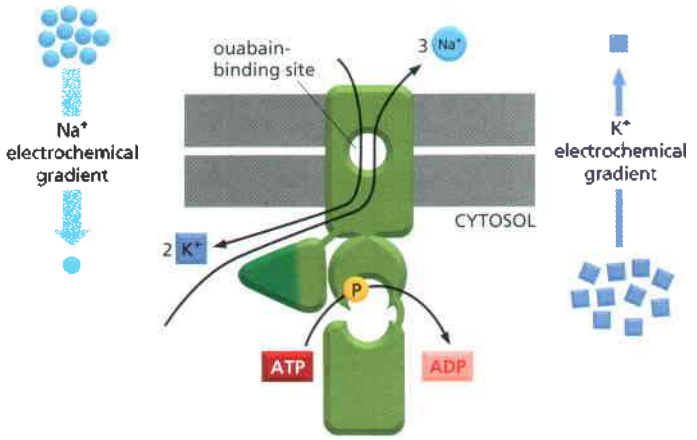


Figure 11–14 The Na^+-K^+ pump. <GAGT> This transporter actively pumps Na^+ out of and K^+ into a cell against their electrochemical gradients. For every molecule of ATP hydrolyzed inside the cell, three Na^+ are pumped out and two K^+ are pumped in. The specific inhibitor ouabain and K^+ compete for the same site on the extracellular side of the pump.

electrically active nerve cells, which, as we shall see, repeatedly gain small amounts of Na^+ and lose small amounts of K^+ during the propagation of nerve impulses.

Like any enzyme, the Na^+-K^+ pump can be driven in reverse, in this case to produce ATP. When the Na^+ and K^+ gradients are experimentally increased to such an extent that the energy stored in their electrochemical gradients is greater than the chemical energy of ATP hydrolysis, these ions move down their electrochemical gradients and ATP is synthesized from ADP and phosphate by the Na^+-K^+ pump. Thus, the phosphorylated form of the pump (step 2 in Figure 11–15) can relax by either donating its phosphate to ADP (step 2 to step 1) or changing its conformation (step 2 to step 3). Whether the overall change in free energy is used to synthesize ATP or to pump Na^+ out of the cell depends on the relative concentrations of ATP, ADP, and phosphate, as well as on the electrochemical gradients for Na^+ and K^+.

Since the Na^+-K^+ pump drives three positively charged ions out of the cell for every two it pumps in, it is *electrogenic*. It drives a net current across the

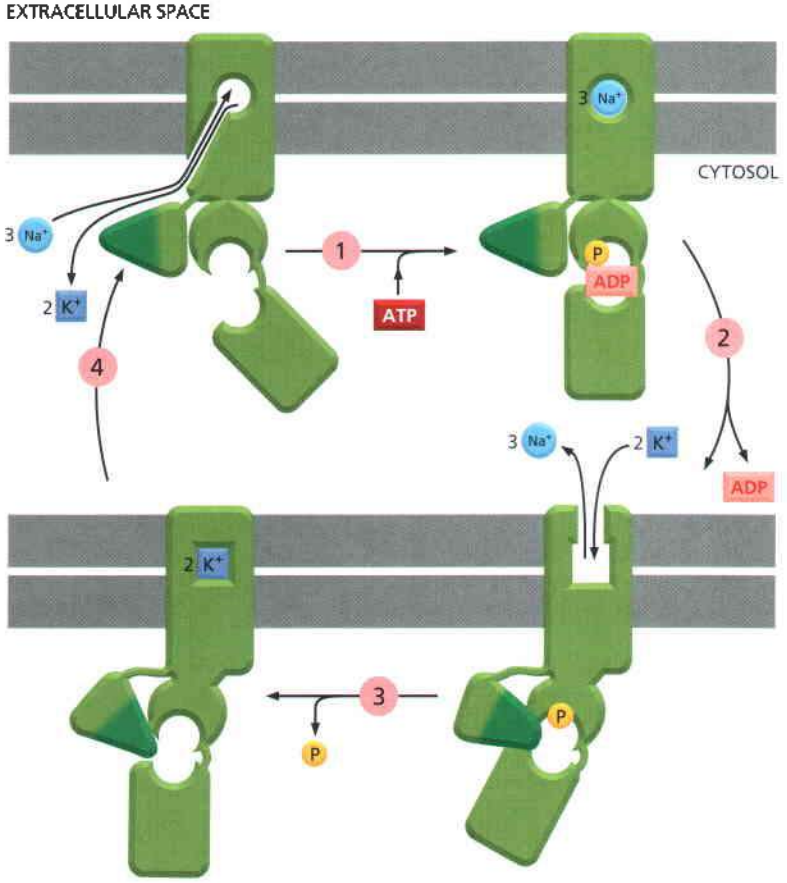


Figure 11–15 A model of the pumping cycle of the Na^+-K^+ pump. (1) The binding of intracellular Na^+ and the subsequent phosphorylation by ATP of the cytoplasmic face of the pump induce a conformational change in the protein that (2) transfers the Na^+ across the membrane and releases it on the outside of the cell. (3) Then, the binding of K^+ on the extracellular surface and the subsequent dephosphorylation of the pump return the protein to its original conformation, which (4) transfers the K^+ across the membrane and releases it into the cytosol. These changes in conformation are analogous to the A ↔ B transitions shown in Figure 11–5, except that here the Na^+-dependent phosphorylation and the K^+-dependent dephosphorylation of the protein cause the conformational transitions to occur in an orderly manner, enabling the protein to do useful work. Although for simplicity the diagram shows only one Na^+- and one K^+-binding site, in the real pump there are three Na^+- and two K^+-binding sites.

membrane, tending to create an electrical potential, with the cell's inside being negative relative to the outside. This electrogenic effect of the pump, however, seldom contributes more than 10% to the membrane potential. The remaining 90%, as we discuss later, depends only indirectly on the Na^+-K^+ pump.

On the other hand, the Na^+-K^+ pump does have a direct role in controlling the solute concentration inside the cell and thereby helps regulate **osmolarity** (or *tonicity*) of the cytosol. All cells contain specialized water channel proteins called *aquaporins* (discussed in detail on p. 673) in their plasma membrane to facilitate water flow across this membrane. Thus, water moves into or out of cells down its concentration gradient, a process called *osmosis*. As explained in **Panel 11–1**, cells contain a high concentration of solutes, including numerous negatively charged organic molecules that are confined inside the cell (the so-called *fixed anions*) and their accompanying cations that are required for charge balance. This tends to create a large osmotic gradient that tends to "pull" water into the cell. Animal cells counteract this effect by an opposite osmotic gradient due to a high concentration of inorganic ions—chiefly Na^+ and Cl^-—in the extracellular fluid. The Na^+-K^+ pump helps maintain osmotic balance by pumping out the Na^+ that leaks in down its steep electrochemical gradient. The Cl^- is kept out by the membrane potential.

In the special case of human red blood cells, which lack a nucleus and other organelles and have a plasma membrane that has an unusually high permeability to water, osmotic water movements can greatly influence cell volume, and the Na^+-K^+ pump plays an important part in maintaining red cell volume. If these cells are placed in a *hypotonic solution* (that is, a solution having a low solute concentration and therefore a high water concentration), there is net movement of water into the cells, causing them to swell and burst (lyse); conversely, if the cells are placed in a *hypertonic solution*, they shrink (**Figure 11–16**). The role of the Na^+-K^+ pump in controlling red cell volume is indicated by the observation that the cells swell, and may eventually burst, if they are treated with *ouabain*, which inhibits the Na^+-K^+ pump. For most animal cells, however, osmosis and the Na^+-K^+ pump have only minor roles in regulating cell volume. This is because most of the cytoplasm is in a gel-like state and resists large changes in its volume in response to changes in osmolarity.

Nonanimal cells cope with their osmotic problems in various ways. Plant cells and many bacteria are prevented from bursting by the semirigid cell wall that surrounds their plasma membrane. In amoebae, the excess water that flows in osmotically is collected in contractile vacuoles, which periodically discharge their contents to the exterior (see Panel 11–1). Bacteria have also evolved strategies that allow them to lose ions, and even macromolecules, quickly when subjected to an osmotic shock.

ABC Transporters Constitute the Largest Family of Membrane Transport Proteins

The last type of carrier protein that we discuss is the family of the **ABC transporters**, so named because each member contains two highly conserved ATPase domains or ATP-binding "cassettes" (**Figure 11–17**). ATP binding leads to dimerization of the two ATP-binding domains, and ATP hydrolysis leads to

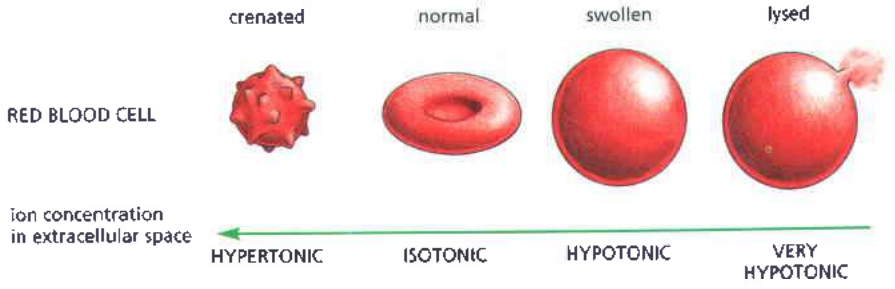
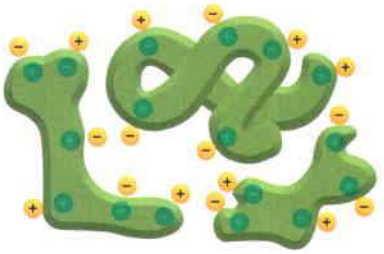
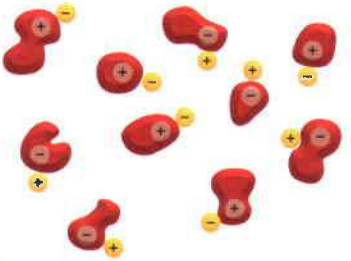


Figure 11–16 Response of a human red blood cell to changes in osmolarity of the extracellular fluid. <GTAC> The cell swells or shrinks as water moves into or out of the cell down its concentration gradient.

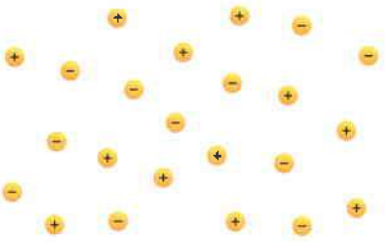
SOURCES OF INTRACELLULAR OSMOLARITY



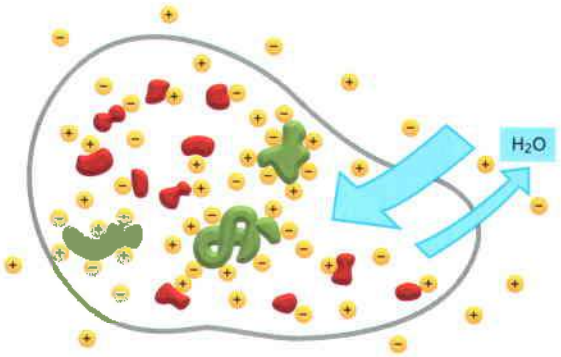
1 Macromolecules themselves contribute very little to the osmolarity of the cell interior since, despite their large size, each one counts only as a single molecule and there are relatively few of them compared to the number of small molecules in the cell.However, most biological macromolecules are highly charged, and they attract many inorganic ions of opposite charge. Because of their large numbers, these counterions make a major contribution to intracellular osmolarity.



2 As the result of active transport and metabolic processes, the cell contains a high concentration of small organic molecules, such as sugars, amino acids, and nucleotides, to which its plasma membrane is impermeable. Because most of these metabolites are charged, they also attract counterions. Both the small metabolites and their counterions make a further major contribution to intracellular osmolarity.



3 The osmolarity of the extracellular fluid is usually due mainly to small inorganic ions. These leak slowly across the plasma membrane into the cell. If they were not pumped out, and if there were no other molecules inside the cell that interacted with them so as to influence their distribution, they would eventually come to equilibrium with equal concentrations inside and outside the cell. However, the presence of charged macromolecules and metabolites in the cell that attract these ions gives rise to the Donnan effect: it causes the total concentration of inorganic ions (and therefore their contribution to the osmolarity) to be greater inside than outside the cell at equilibrium.

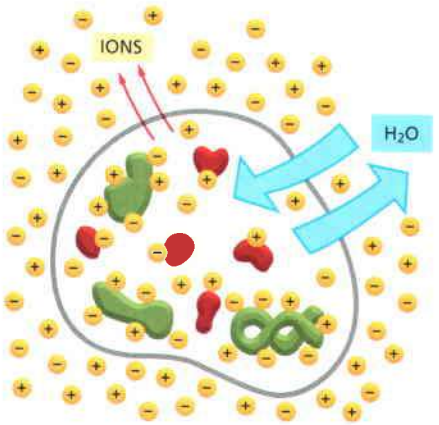


THE PROBLEM

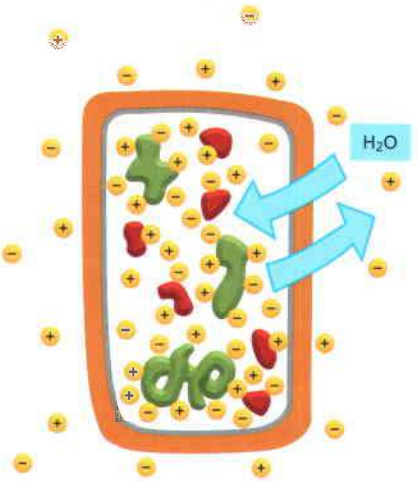
Because of the above factors, a cell that does nothing to control its osmolarity will have a higher concentration of solutes inside than outside. As a result, water will be higher in concentration outside the cell than inside. This difference in water concentration across the plasma membrane will cause water to move continuously into the cell by osmosis.

THE SOLUTION

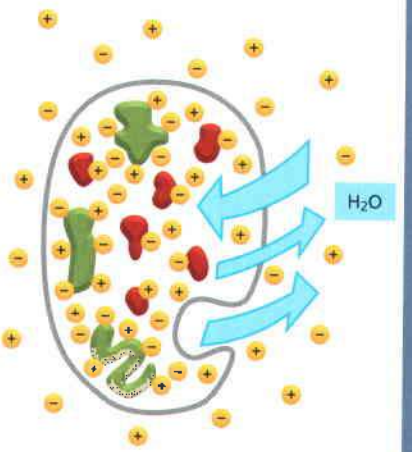
Animal cells and bacteria control their intracellular osmolarity by actively pumping out inorganic ions, such as Na^+, so that their cytoplasm contains a lower total concentration of inorganic ions than the extracellular fluid, thereby compensating for their excess of organic solutes.



Plant cells are prevented from swelling by their rigid walls and so can tolerate an osmotic difference across their plasma membranes: an internal turgor pressure is built up, which at equilibrium forces out as much water as enters.



Many protozoa avoid becoming swollen with water, despite an osmotic difference across the plasma membrane, by periodically extruding water from special contractile vacuoles.



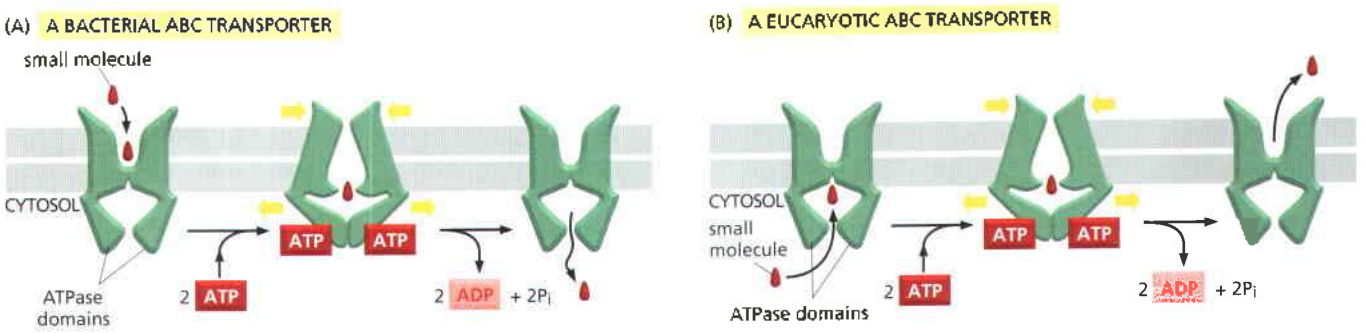


Figure 11–17 Typical ABC transporters in procaryotes (A) and eucaryotes (B). Transporters consist of multiple domains: typically, two hydrophobic domains, each built of six membrane-spanning segments that form the translocation pathway and provide substrate specificity, and two ATPase domains (also called ATP-binding cassettes) protruding into the cytosol. In some cases, the two halves of the transporter are formed by a single polypeptide, whereas in other cases they are formed by two or more separate polypeptides that assemble into a similar structure (see Figure 10–24). Without ATP bound, the transporter exposes a substrate-binding site to either the extracellular space (in procaryotes) or the intracellular space (in eucaryotes or procaryotes). ATP binding induces a conformational change that exposes the substrate-binding pocket to the opposite face; ATP hydrolysis followed by ADP dissociation returns the transporter to its original conformation. Most individual ABC transporters are unidirectional. Both importing and exporting ABC transporters are found in bacteria, but in eucaryotes almost all ABC transporters export substances from the cytosol—either to the extracellular space or to a membrane-bound intracellular compartment such as the ER or the mitochondria.

their dissociation. These structural changes in the cytosolic domains are thought to be transmitted to the transmembrane segments, driving cycles of conformational changes that alternately expose substrate-binding sites on one side of the membrane and then on the other. In this way, ABC transporters use ATP binding and hydrolysis to transport small molecules across the bilayer.

ABC transporters constitute the largest family of membrane transport proteins and are of great clinical importance. The first of these proteins to be characterized was found in bacteria. We have already mentioned that the plasma membranes of all bacteria contain transporters that use the H^+ gradient across the membrane to pump a variety of nutrients into the cell. Bacteria also have transport ATPases that use the energy of ATP hydrolysis to import certain small molecules. In bacteria such as *E. coli*, which have double membranes (**Figure 11–18**), the transport ATPases are located in the inner membrane, and an auxiliary mechanism operates to capture the nutrients and deliver them to the transporters (**Figure 11–19**).

In *E. coli*, 78 genes (an amazing 5% of the bacterium's genes) encode ABC transporters, and animal genomes encode more. Although each transporter is thought to be specific for a particular molecule or class of molecules, the variety of substrates transported by this superfamily is great and includes inorganic ions, amino acids, mono- and polysaccharides, peptides, and even proteins. Whereas bacterial ABC transporters are used for both import and export, those identified in eucaryotes seem mostly specialized for export.

Indeed, the first eucaryotic ABC transporters identified were discovered because of their ability to pump hydrophobic drugs out of the cytosol. One of these transporters is the **multidrug resistance (MDR) protein**, the overexpression

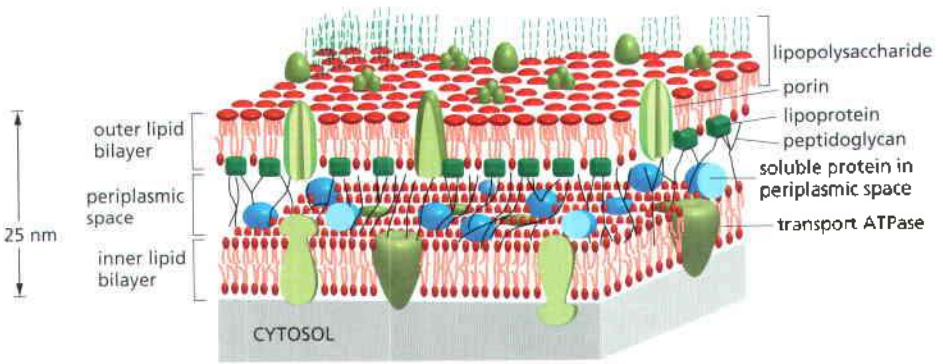


Figure 11–18 A small section of the double membrane of an *E. coli* bacterium. The inner membrane is the cell's plasma membrane. Between the inner and outer lipid bilayer membranes is a highly porous, rigid peptidoglycan layer, composed of protein and polysaccharide, that constitutes the bacterial cell wall. It is attached to lipoprotein molecules in the outer membrane and fills the *periplasmic space* (only a little of the peptidoglycan layer is shown). This space also contains a variety of soluble protein molecules. The dashed threads (shown in *green*) at the top represent the polysaccharide chains of the special lipopolysaccharide molecules that form the external monolayer of the outer membrane; for clarity, only a few of these chains are shown. Bacteria with double membranes are called *Gram-negative* because they do not retain the dark blue dye used in Gram staining. Bacteria with single membranes (but thicker cell walls), such as staphylococci and streptococci, retain the blue dye and are therefore called *Gram-positive*; their single membrane is analogous to the inner (plasma) membrane of Gram-negative bacteria.

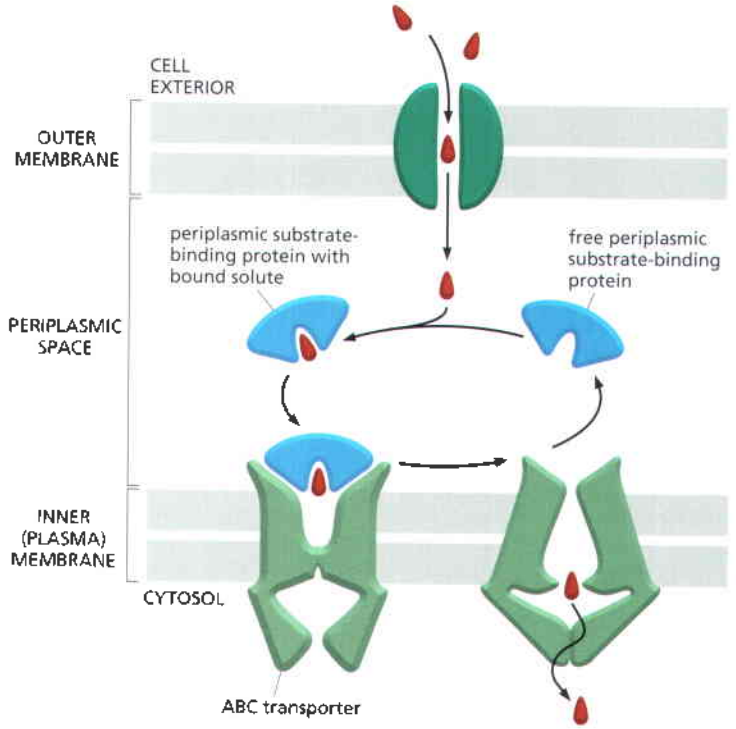


Figure 11–19 The auxiliary transport system associated with transport ATPases in bacteria with double membranes. The solute diffuses through channel-forming proteins (porins) in the outer membrane and binds to a *periplasmic substrate-binding protein*, which undergoes a conformational change that enables it to bind to an ABC transporter in the plasma membrane. The ABC transporter then picks up the solute and actively transfers it across the plasma membrane in a reaction driven by ATP hydrolysis. The peptidoglycan is omitted for simplicity; its porous structure allows the substrate-binding proteins and water-soluble solutes to move through it by simple diffusion.

of which in human cancer cells can make the cells simultaneously resistant to a variety of chemically unrelated cytotoxic drugs that are widely used in cancer chemotherapy. Treatment with any one of these drugs can result in the selective survival and overgrowth of those cancer cells that express more of the MDR transporter. These cells can pump drugs out of the cell very efficiently and are therefore relatively resistant to the toxic effects of the anticancer drugs. Selection for cancer cells with resistance to one drug can thereby lead to resistance to a wide variety of anti-cancer drugs. Some studies indicate that up to 40% of human cancers develop multidrug resistance, making it a major hurdle in the battle against cancer.

A related and equally sinister phenomenon occurs in the protist *Plasmodium falciparum*, which causes malaria. More than 200 million people are infected worldwide with this parasite, which remains a major cause of human death, killing more than a million people every year. The development of resistance to the antimalarial drug *chloroquine* hampers the control of malaria. The resistant *P. falciparum* have amplified a gene encoding an ABC transporter that pumps out the chloroquine.

In most vertebrate cells, an ABC transporter in the endoplasmic reticulum (ER) membrane actively transports a wide variety of peptides from the cytosol into the ER lumen. These peptides are produced by protein degradation in proteasomes (discussed in Chapter 6). They are carried from the ER to the cell surface, where they are displayed for scrutiny by cytotoxic T lymphocytes, which will kill the cell if the peptides are derived from a virus or other microorganisms lurking in the cytosol of an infected cell (discussed in Chapter 25).

Yet another member of the ABC transporter family is the *cystic fibrosis transmembrane conductance regulator* protein (CFTR), which was discovered through studies of the common genetic disease *cystic fibrosis*. This disease is caused by a mutation in the gene encoding CFTR, which functions as a Cl^- channel in the plasma membrane of epithelial cells. CFTR regulates ion concentrations in the extracellular fluid, especially in the lung. One in 27 Caucasians carries a gene encoding a mutant form of this protein; in 1 in 2900, both copies of the gene are mutated, causing the disease. In contrast to other ABC transporters, ATP binding and hydrolysis do not drive the transport process. Instead, they control the opening and closing of the Cl^- channel, which provides a passive conduit for Cl^- to move down its electrochemical gradient. Thus, ABC proteins can apparently function as either transporters or channels.

Summary

Transporters bind specific solutes and transfer them across the lipid bilayer by undergoing conformational changes that expose the solute-binding site sequentially on one side of the membrane and then on the other. Some transporters simply move a single solute "downhill," whereas others can act as pumps to move a solute "uphill" against its electrochemical gradient, using energy provided by ATP hydrolysis, by a downhill flow of another solute (such as Na^+ or H^+), or by light to drive the requisite series of conformational changes in an orderly manner. Transporters belong to a small number of protein families. Each family contains proteins of similar amino acid sequences that are thought to have evolved from a common ancestral protein and to operate by a similar mechanism. The family of P-type transport ATPases, which includes Ca^{2+} and Na^+-K^+ pumps, is an important example; each of these ATPases sequentially phosphorylates and dephosphorylates itself during the pumping cycle. The superfamily of ABC transporters is the largest family of membrane transport proteins and is especially important clinically. It includes proteins that are responsible for cystic fibrosis, as well as for drug resistance in both cancer cells and malaria-causing parasites.

ION CHANNELS AND THE ELECTRICAL PROPERTIES OF MEMBRANES

Unlike carrier proteins, channel proteins form hydrophilic pores across membranes. One class of channel proteins found in virtually all animals forms *gap junctions* between two adjacent cells; each plasma membrane contributes equally to the formation of the channel, which connects the cytoplasm of the two cells. These channels are discussed in Chapter 19 and will not be considered further here. Both gap junctions and *porins*, the channel-forming proteins of the outer membranes of bacteria, mitochondria, and chloroplasts (discussed in Chapter 10), have relatively large and permissive pores, which would be disastrous if they directly connected the inside of a cell to an extracellular space. Indeed, many bacterial toxins do exactly that to kill other cells (discussed in Chapter 24).

In contrast, most channel proteins in the plasma membrane of animal and plant cells that connect the cytosol to the cell exterior necessarily have narrow, highly selective pores that can open and close rapidly. Because these proteins are concerned specifically with inorganic ion transport, they are referred to as **ion channels**. For transport efficiency, ion channels have an advantage over carriers in that up to 100 million ions can pass through one open channel each second—a rate 10^5 times greater than the fastest rate of transport mediated by any known carrier protein. However, channels cannot be coupled to an energy source to perform active transport, so the transport that they mediate is always passive (downhill). Thus, the function of ion channels is to allow specific inorganic ions—primarily Na^+, K^+, Ca^{2+}, or Cl^-—to diffuse rapidly down their electrochemical gradients across the lipid bilayer. As we shall see, the ability to control ion fluxes through these channels is essential for many cell functions. Nerve cells (neurons), in particular, have made a specialty of using ion channels, and we shall consider how they use many different ion channels to receive, conduct, and transmit signals.

Ion Channels Are Ion-Selective and Fluctuate Between Open and Closed States

Two important properties distinguish ion channels from simple aqueous pores. First, they show *ion selectivity*, permitting some inorganic ions to pass, but not others. This suggests that their pores must be narrow enough in places to force permeating ions into intimate contact with the walls of the channel so that only ions of appropriate size and charge can pass. The permeating ions have to shed most or all of their associated water molecules to pass, often in single file,

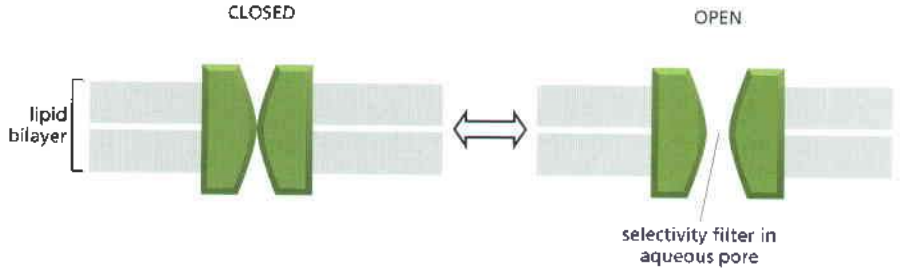


Figure 11–20 **A typical ion channel, which fluctuates between closed and open conformations.** The channel protein shown here in cross section forms a hydrophilic pore across the lipid bilayer only in the "open" conformational state. Polar groups are thought to line the wall of the pore, while hydrophobic amino acid side chains interact with the lipid bilayer (not shown). The pore narrows to atomic dimensions in one region (the selectivity filter), where the ion selectivity of the channel is largely determined.

through the narrowest part of the channel, which is called the *selectivity filter*; this limits their rate of passage (**Figure 11–20**). Thus, as the ion concentration increases, the flux of the ion through a channel increases proportionally but then levels off (saturates) at a maximum rate.

The second important distinction between ion channels and simple aqueous pores is that ion channels are not continuously open. Instead, they are *gated*, which allows them to open briefly and then close again (**Figure 11–21**). Moreover, with prolonged (chemical or electrical) stimulation, most channels go into a closed "desensitized" or "inactivated" state, in which they are refractory to further opening until the stimulus has been removed, as we discuss later. In most cases, the gate opens in response to a specific stimulus. The main types of stimuli that are known to cause ion channels to open are a change in the voltage across the membrane *(voltage-gated channels)*, a mechanical stress *(mechanically gated channels)*, or the binding of a ligand *(ligand-gated channels)*. The ligand can be either an extracellular mediator—specifically, a neurotransmitter *(transmitter-gated channels)*—or an intracellular mediator such as an ion *(ion-gated channels)* or a nucleotide *(nucleotide-gated channels)*. In addition, protein phosphorylation and dephosphorylation regulates the activity of many ion channels; this type of channel regulation is discussed, together with nucleotide-gated ion channels, in Chapter 15.

More than 100 types of ion channels have been described thus far, and new ones are still being discovered, each characterized by the ions it conducts, the mechanism by which it is gated, and its abundance and localization in the cell. Ion channels are responsible for the electrical excitability of muscle cells, and they mediate most forms of electrical signaling in the nervous system. A single neuron might typically contain 10 or more kinds of ion channels, located in different domains of its plasma membrane. But ion channels are not restricted to electrically excitable cells. They are present in all animal cells and are found in plant cells and microorganisms: they propagate the leaf-closing response of the mimosa plant, for example, and allow the single-celled *Paramecium* to reverse direction after a collision.

Perhaps the most common ion channels are those that are permeable mainly to K^+. These channels are found in the plasma membrane of almost all

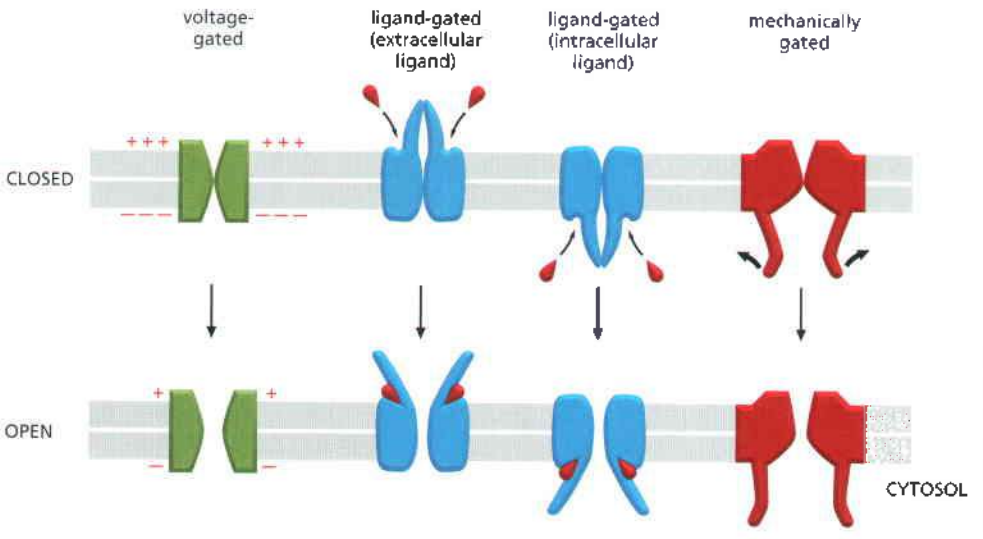


Figure 11–21 **The gating of ion channels.** This drawing shows different kinds of stimuli that open ion channels. Mechanically gated channels often have cytoplasmic extensions that link the channel to the cytoskeleton (not shown).

animal cells. An important subset of K^+ channels opens even in an unstimulated or "resting" cell, and hence these channels are sometimes called **K^+ leak channels**. Although this term applies to many different K^+ channels, depending on the cell type, they serve a common purpose. By making the plasma membrane much more permeable to K^+ than to other ions, they have a crucial role in maintaining the membrane potential across all plasma membranes.

The Membrane Potential in Animal Cells Depends Mainly on K^+ Leak Channels and the K^+ Gradient Across the Plasma Membrane

A **membrane potential** arises when there is a difference in the electrical charge on the two sides of a membrane, due to a slight excess of positive ions over negative ones on one side and a slight deficit on the other. Such charge differences can result both from active electrogenic pumping (see p. 662) and from passive ion diffusion. As we discuss in Chapter 14, electrogenic H^+ pumps in the mitochondrial inner membrane generate most of the membrane potential across this membrane. Electrogenic pumps also generate most of the electrical potential across the plasma membrane in plants and fungi. In typical animal cells, however, passive ion movements make the largest contribution to the electrical potential across the plasma membrane.

As explained earlier, the Na^+-K^+ pump helps maintain an osmotic balance across the animal cell membrane by keeping the intracellular concentration of Na^+ low. Because there is little Na^+ inside the cell, other cations have to be plentiful there to balance the charge carried by the cell's fixed anions—the negatively charged organic molecules that are confined inside the cell. This balancing role is performed largely by K^+, which is actively pumped into the cell by the Na^+-K^+ pump and can also move freely in or out through the *K^+ leak channels* in the plasma membrane. Because of the presence of these channels, K^+ comes almost to equilibrium, where an electrical force exerted by an excess of negative charges attracting K^+ into the cell balances the tendency of K^+ to leak out down its concentration gradient. The membrane potential is the manifestation of this electrical force, and we can calculate its equilibrium value from the steepness of the K^+ concentration gradient. The following argument may help to make this clear.

Suppose that initially there is no voltage gradient across the plasma membrane (the membrane potential is zero) but the concentration of K^+ is high inside the cell and low outside. K^+ will tend to leave the cell through the K^+ leak channels, driven by its concentration gradient. As K^+ begins to move out, each ion leaves behind an unbalanced negative charge, thereby creating an electrical field, or membrane potential, which will tend to oppose the further efflux of K^+. The net efflux of K^+ halts when the membrane potential reaches a value at which this electrical driving force on K^+ exactly balances the effect of its concentration gradient—that is, when the electrochemical gradient for K^+ is zero. Although Cl^- ions also equilibrate across the membrane, the membrane potential keeps most of these ions out of the cell because their charge is negative.

The equilibrium condition, in which there is no net flow of ions across the plasma membrane, defines the **resting membrane potential** for this idealized cell. A simple but very important formula, the **Nernst equation**, quantifies the equilibrium condition and, as explained in **Panel 11–2**, makes it possible to calculate the theoretical resting membrane potential if we know the ratio of internal and external ion concentrations. As the plasma membrane of a real cell is not exclusively permeable to K^+ and Cl^-, however, the actual resting membrane potential is usually not exactly equal to that predicted by the Nernst equation for K^+ or Cl^-.

The Resting Potential Decays Only Slowly When the Na^+-K^+ Pump Is Stopped

Only a minute number of ions must move across the plasma membrane to set up the membrane potential. Thus, we can think of the membrane potential as arising from movements of charge that leave ion *concentrations* practically unaffected

THE NERNST EQUATION AND ION FLOW

The flow of any ion through a membrane channel protein is driven by the electrochemical gradient for that ion. This gradient represents the combination of two influences: the voltage gradient and the concentration gradient of the ion across the membrane. When these two influences just balance each other the electrochemical gradient for the ion is zero and there is no *net* flow of the ion through the channel. The voltage gradient (membrane potential) at which this equilibrium is reached is called the equilibrium potential for the ion. It can be calculated from an equation that will be derived below, called the Nernst equation.

The Nernst equation is

$$V = \frac{RT}{zF} \ln \frac{C_o}{C_i}$$

where

V = the equilibrium potential in volts (internal potential minus external potential)
C_o and C_i = outside and inside concentrations of the ion, respectively
R = the gas constant (2 cal mol^{-1} K^{-1})
T = the absolute temperature (K)
F = Faraday's constant (2.3×10^4 cal V^{-1} mol^{-1})
z = the valence (charge) of the ion
ln = logarithm to the base e

The Nernst equation is derived as follows:

A molecule in solution (a solute) tends to move from a region of high concentration to a region of low concentration simply due to the random movement of molecules, which results in their equilibrium. Consequently, movement down a concentration gradient is accompanied by a favorable free-energy change ($\Delta G < 0$), whereas movement up a concentration gradient is accompanied by an unfavorable free-energy change ($\Delta G > 0$). (Free energy is introduced in Chapter 2, and discussed in the context of redox reactions in Panel 14–1, p. 830.)

The free-energy change per mole of solute moved across the plasma membrane (ΔG_{conc}) is equal to $-RT \ln C_o / C_i$.

If the solute is an ion, moving it into a cell across a membrane whose inside is at a voltage V relative to the outside will cause an additional free-energy change (per mole of solute moved) of $\Delta G_{volt} = zFV$.

At the point where the concentration and voltage gradients just balance,

$$\Delta G_{conc} + \Delta G_{volt} = 0$$

and the ion distribution is at equilibrium across the membrane.

Thus,

$$zFV - RT \ln \frac{C_o}{C_i} = 0$$

and, therefore,

$$V = \frac{RT}{zF} \ln \frac{C_o}{C_i}$$

or, using the constant that converts natural logarithms to base 10,

$$V = 2.3 \frac{RT}{zF} \log_{10} \frac{C_o}{C_i}$$

For a univalent ion,

$$2.3 \frac{RT}{F} = 58 \text{ mV at } 20^\circ C \quad \text{and} \quad 61.5 \text{ mV at } 37^\circ C$$

Thus, for such an ion at 37°C,

$$V = +61.5 \text{ mV for } C_o / C_i = 10,$$

whereas

$$V = 0 \text{ for } C_o / C_i = 1.$$

The K^+ equilibrium potential (V_K), for example, is

$$61.5 \log_{10}([K^+]_o / [K^+]_i) \text{ millivolts}$$

(–89 mV for a typical cell where $[K^+]_o$ = 5 mM and $[K^+]_i$ = 140 mM).

At V_K, there is no net flow of K^+ across the membrane.

Similarly, when the membrane potential has a value of

$$61.5 \log_{10}([Na^+]_o / [Na^+]_i),$$

the Na^+ equilibrium potential (V_{Na}),

there is no net flow of Na^+.

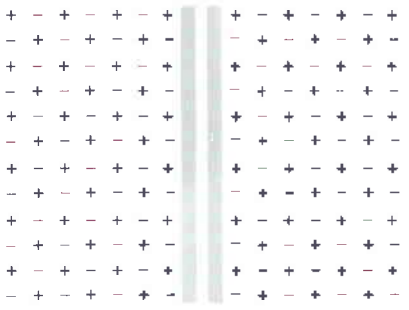
For any particular membrane potential, V_M, the net force tending to drive a particular type of ion out of the cell, is proportional to the difference between V_M and the equilibrium potential for the ion: hence,

for K^+ it is $V_M - V_K$

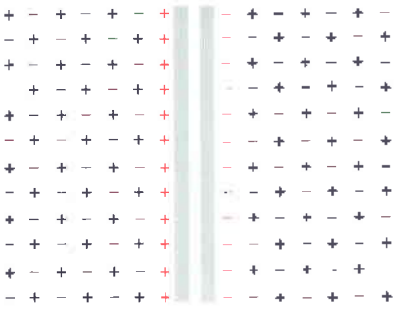
and for Na^+ it is $V_M - V_{Na}$.

The number of ions that go to form the layer of charge adjacent to the membrane is minute compared with the total number inside the cell. For example, the movement of 6000 Na^+ ions across 1 μm^2 of membrane will carry sufficient charge to shift the membrane potential by about 100 mV.

Because there are about 3×10^7 Na^+ ions in a typical cell (1 μm^3 of bulk cytoplasm), such a movement of charge will generally have a negligible effect on the ion concentration gradients across the membrane.



exact balance of charges on each side of the membrane; membrane potential = 0



a few of the positive ions *(red)* cross the membrane from right to left, leaving their negative counterions *(red)* behind; this sets up a nonzero membrane potential

Figure 11–22 The ionic basis of a membrane potential. A small flow of ions carries sufficient charge to cause a large change in the membrane potential. The ions that give rise to the membrane potential lie in a thin (< 1 nm) surface layer close to the membrane, held there by their electrical attraction to their oppositely charged counterparts (counterions) on the other side of the membrane. For a typical cell, 1 microcoulomb of charge (6×10^{12} monovalent ions) per square centimeter of membrane, transferred from one side of the membrane to the other, changes the membrane potential by roughly 1 V. This means, for example, that in a spherical cell of diameter 10 μm, the number of K^+ ions that have to flow out to alter the membrane potential by 100 mV is only about 1/100,000 of the total number of K^+ ions in the cytosol.

and result in only a very slight discrepancy in the number of positive and negative ions on the two sides of the membrane (**Figure 11–22**). Moreover, these movements of charge are generally rapid, taking only a few milliseconds or less.

Consider the change in the membrane potential in a real cell after the sudden inactivation of the Na^+-K^+ pump. A slight drop in the membrane potential occurs immediately. This is because the pump is electrogenic and, when active, makes a small direct contribution to the membrane potential by pumping out three Na^+ for every two K^+ that it pumps in. However, switching off the pump does not abolish the major component of the resting potential, which is generated by the K^+ equilibrium mechanism described above. This component of the membrane potential persists as long as the Na^+ concentration inside the cell stays low and the K^+ ion concentration high—typically for many minutes. But the plasma membrane is somewhat permeable to all small ions, including Na^+. Therefore, without the Na^+-K^+ pump, the ion gradients set up by the pump will eventually run down, and the membrane potential established by diffusion through the K^+ leak channels will fall as well. As Na^+ enters, the osmotic balance is upset, and water seeps into the cell (see Panel 11–1, p. 664), and the cell eventually comes to a new resting state where Na^+, K^+, and Cl^- are all at equilibrium across the membrane. The membrane potential in this state is much less than it was in the normal cell with an active Na^+-K^+ pump.

The resting potential of an animal cell varies between –20 mV and –120 mV, depending on the organism and cell type. Although the K^+ gradient always has a major influence on this potential, the gradients of other ions (and the disequilibrating effects of ion pumps) also have a significant effect: the more permeable the membrane for a given ion, the more strongly the membrane potential tends to be driven toward the equilibrium value for that ion. Consequently, changes in a membrane's permeability to ions can cause significant changes in the membrane potential. This is one of the key principles relating the electrical excitability of cells to the activities of ion channels.

To understand how ion channels select their ions and how they open and close, we need to know their atomic structure. The first ion channel to be crystallized and studied by x-ray diffraction was a bacterial K^+ channel. The details of its structure revolutionized our understanding of ion channels.

The Three-Dimensional Structure of a Bacterial K^+ Channel Shows How an Ion Channel Can Work

Scientists were puzzled by the remarkable ability of ion channels to combine exquisite ion selectivity with a high conductance. K^+ leak channels, for example, conduct K^+ 10,000-fold better than Na^+, yet the two ions are both featureless spheres and have similar diameters (0.133 nm and 0.095 nm, respectively). A single amino acid substitution in the pore of an animal cell K^+ channel can result in a loss of ion selectivity and cell death. We cannot explain the normal K^+ selectivity by pore size, because Na^+ is smaller than K^+. Moreover, the high conductance rate is incompatible with the channel's having selective, high-affinity K^+-binding sites, as the binding of K^+ ions to such sites would greatly slow their passage.

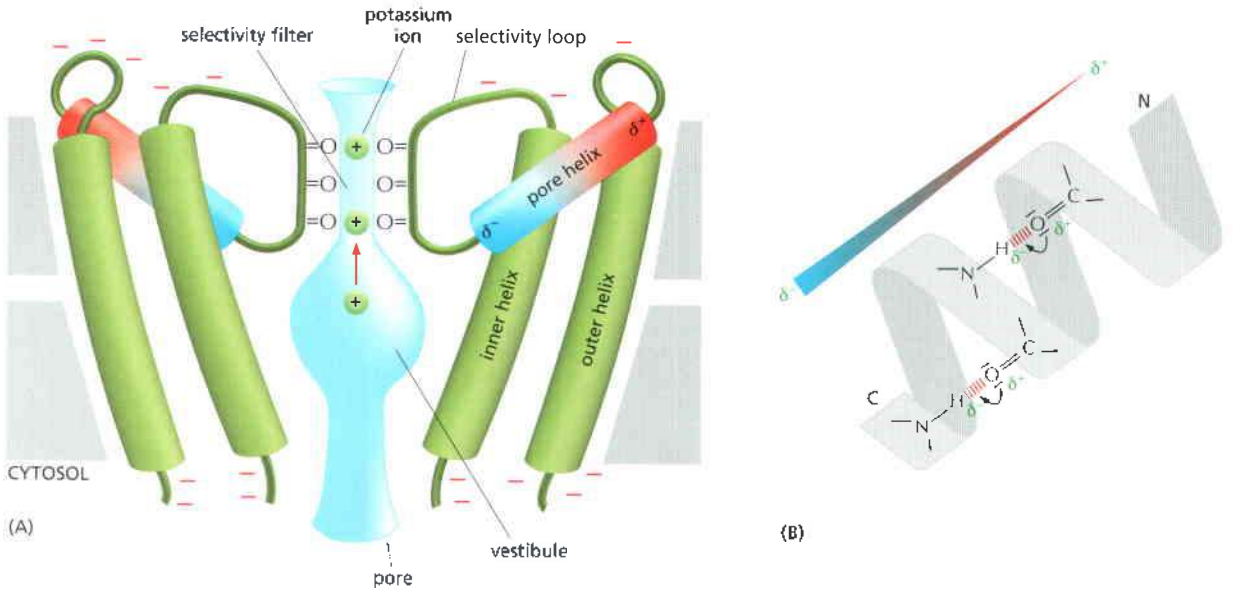


Figure 11–23 **The structure of a bacterial K^+ channel.** <ATTA> (A) Two transmembrane α helices from only two of the four identical subunits are shown. From the cytosolic side, the pore opens up into a vestibule in the middle of the membrane. The vestibule facilitates transport by allowing the K^+ ions to remain hydrated even though they are halfway across the membrane. The narrow selectivity filter links the vestibule to the outside of the cell. Carbonyl oxygens line the walls of the selectivity filter and form transient binding sites for dehydrated K^+ ions. The positions of the K^+ ions in the pore were determined by soaking crystals of the channel protein in a solution containing rubidium ions, which are more electron-dense but only slightly larger than K^+ ions; from the differences in the diffraction patterns obtained with K^+ ions and with rubidium ions in the channel, the positions of the ions could be calculated. Two K^+ ions occupy sites in the selectivity filter, while a third K^+ ion is located in the center of the vestibule, where it is stabilized by electrical interactions with the more negatively charged ends of the pore helices. The ends of the four pore helices (only two of which are shown) point precisely toward the center of the vestibule, thereby guiding K^+ ions into the selectivity filter. Negatively charged amino acids (indicated by *red* minus signs) are concentrated near the channel entrance and exit. (B) Because of the polarity of the hydrogen bonds (*red*) that link adjacent turns of an α helix, every α helix has an electric dipole along its axis, with a more negatively charged C-terminal end (δ^-) and a more positively charged N-terminal end (δ^+). (A, adapted from D.A. Doyle et al., *Science* 280:69–77, 1998. With permission from AAAS.)

The puzzle was solved when the structure of a *bacterial K^+ channel* was determined by x-ray crystallography. The channel is made from four identical transmembrane subunits, which together form a central pore through the membrane (**Figure 11–23**). Negatively charged amino acids concentrated at the cytosolic entrance to the pore are thought to attract cations and repel anions, making the channel cation-selective. Each subunit contributes two transmembrane α helices, which are tilted outward in the membrane and together form a cone, with its wide end facing the outside of the cell where K^+ ions exit from the channel. The polypeptide chain that connects the two transmembrane helices forms a short α helix (the *pore helix*) and a crucial loop that protrudes into the wide section of the cone to form the **selectivity filter.** The selectivity loops from the four subunits form a short, rigid, narrow pore, which is lined by the carbonyl oxygen atoms of their polypeptide backbones. Because the selectivity loops of all known K^+ channels have similar amino acid sequences, it is likely that they form a closely similar structure. The crystal structure shows two K^+ ions in single file within the selectivity filter, separated by about 0.8 nm. Mutual repulsion between the two ions is thought to help move them through the pore into the extracellular fluid.

The structure of the selectivity filter explains the ion selectivity of the channel. A K^+ ion must lose almost all of its bound water molecules to enter the filter, where it interacts instead with the carbonyl oxygens lining the filter; the oxygens are rigidly spaced at the exact distance to accommodate a K^+ ion. A Na^+ ion, in contrast, cannot enter the filter because the carbonyl oxygens are too far away from the smaller Na^+ ion to compensate for the energy expense associated with the loss of water molecules required for entry (**Figure 11–24**).

Structural studies of K^+ channels and other channels have also indicated some general principles of how channels may open and close. This gating seems

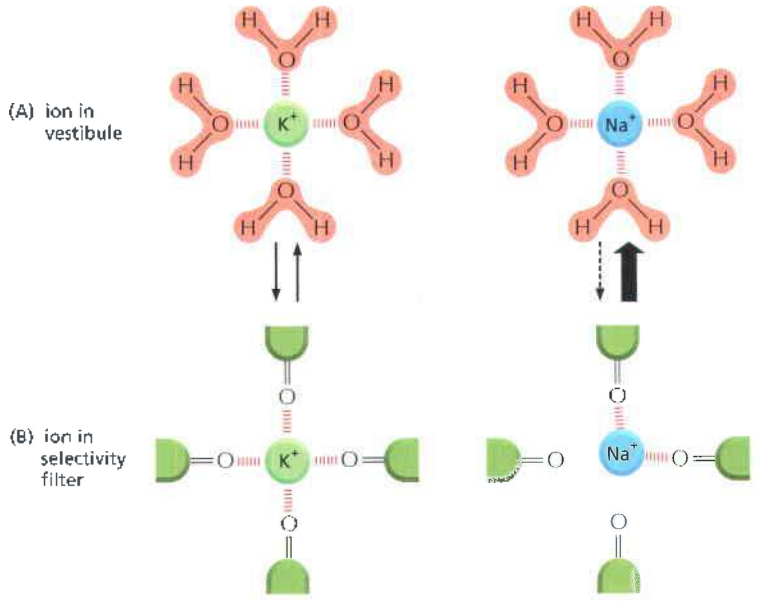


Figure 11–24 K^+ specificity of the selectivity filter in a K^+ channel. The drawing shows K^+ and Na^+ ions (A) in the vestibule and (B) in the selectivity filter of the pore, viewed in cross section. In the vestibule, the ions are hydrated. In the selectivity filter, they have lost their water, and the carbonyl oxygens are placed precisely to accommodate a dehydrated K^+ ion. The dehydration of the K^+ ion requires energy, which is precisely balanced by the energy regained by the interaction of the ion with the carbonyl oxygens that serve as surrogate water molecules. Because the Na^+ ion is too small to interact with the oxygens, it can enter the selectivity filter only at a great energetic expense. The filter therefore selects K^+ ions with high specificity. (Adapted from D.A. Doyle et al., *Science* 280:69–77, 1998. With permission from AAAS.)

to involve movement of the helices in the membrane so that they either obstruct (in the closed state) or free (in the open state) the path for ion movement. Depending on the particular type of channel, helices are thought to tilt, rotate, or bend during gating. The structure of a closed K^+ channel shows that by tilting the inner helices, the pore constricts like a diaphragm at its cytosolic end (Figure 11–25). Bulky hydrophobic amino acid side chains block the small opening that remains, preventing the entry of ions.

Most ion channels are constructed from multiple identical subunits, each of which contributes to a common central pore. A recently determined crystal structure of a Cl^- channel, however, has revealed that some ion channels are built very differently. Although the protein is a dimer formed by two identical subunits, each of the subunits contains its own pore through which Cl^- ions move. In the center of the membrane, amino acid side chains form a selectivity filter, which is conceptually similar to that in K^+ channels. But, unlike the filter in K^+ channels, different regions of the protein contribute the side chains, and they are not symmetrically arranged (Figure 11–26).

Aquaporins Are Permeable to Water But Impermeable to Ions

We discussed earlier that procaryotic and eucaryotic cells have **water channels**, or **aquaporins**, embedded in their plasma membrane to allow water to move readily across this membrane. Aquaporins are especially abundant in cells that must transport water at particularly high rates, such as the epithelial cells of the kidney.

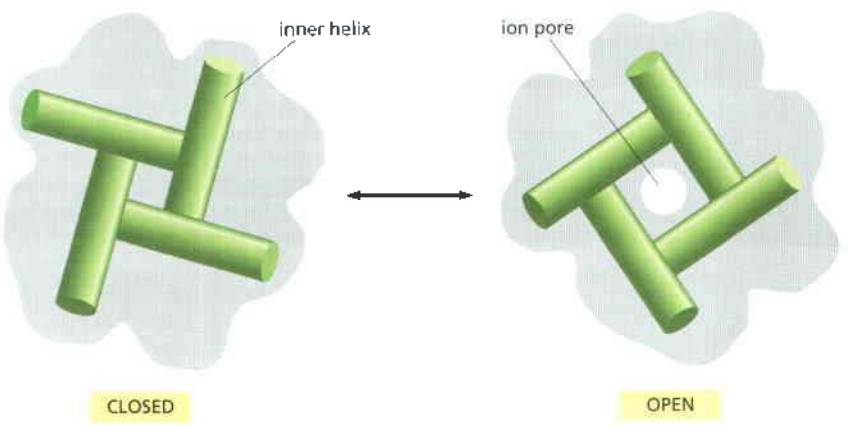


Figure 11–25 A model for the gating of a bacterial K^+ channel. The channel is viewed in cross section. To adopt the closed conformation, the four inner transmembrane helices that line the pore on the cytosolic side of the selectivity filter (see Figure 11–23) rearrange to close the cytosolic entrance to the channel. (Adapted from E. Perozo et al., *Science* 285:73–78, 1999. With permission from AAAS.)

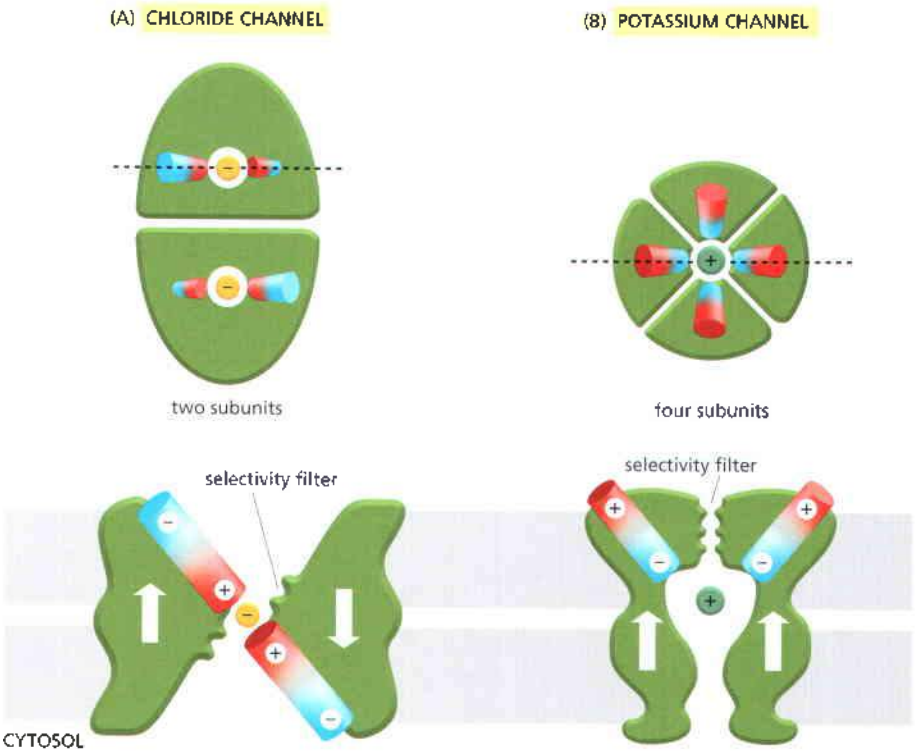


Figure 11–26 Comparison of Cl^- and K^+ channel architectures. (A) The Cl^- channel is a "double-channel" dimer built of two identical subunits, each of which contains its own ion-conducting pore. The upper cartoon is a schematic view of the extracellular face of the channel, showing the two identical ion-conducting pores. The lower cartoon shows a cross section through one subunit, viewed from within the membrane. (*Dotted black line* in top cartoon indicates the plane of section.) The subunit is a single polypeptide chain consisting of two portions, which, though similar, span the membrane with opposite orientations *(white arrows)*: each portion contributes one pore helix oriented such that its positively charged end points towards a centrally positioned selectivity filter. Both elements, the selectivity filter and the helix dipoles, contribute to the selectivity of the channel for negatively charged Cl^- ions. (B) By contrast, the K^+ channel is a tetramer built of four identical subunits, each of which contributes to a centrally located pore. All four subunits have the same orientation in the membrane *(white arrows)*. Four pore helices, one contributed by each subunit, point their negatively charged ends towards a vestibule, stabilizing a positively charged K^+ ion there (also see Figure 11–23). (Lower cartoons in A and B, adapted from R. Dutzler et al., *Nature* 415:287–294, 2002. With permission from Macmillan Publishers Ltd.)

Aquaporins must solve a problem that is opposite to that facing ion channels. To avoid disrupting ion gradients across membranes, they have to allow the rapid passage of water molecules while completely blocking the passage of ions. The crystal structure of an aquaporin reveals how it achieves this remarkable selectivity. The channels have a narrow pore that allows water molecules to traverse the membrane in single file, following the path of carbonyl oxygens that line one side of the pore (**Figure 11–27**A and B). Hydrophobic amino acids line the other side of the pore. The pore is too narrow for any hydrated ion to enter, and the energy cost of dehydrating an ion would be enormous because the hydrophobic wall of the pore cannot interact with a dehydrated ion to compensate for the loss of water. This design readily explains why the aquaporins cannot conduct K^+, Na^+, Ca^{2+}, or Cl^- ions. To understand why these channels are also impermeable to H^+, recall that most protons are present in cells as H_3O^+,

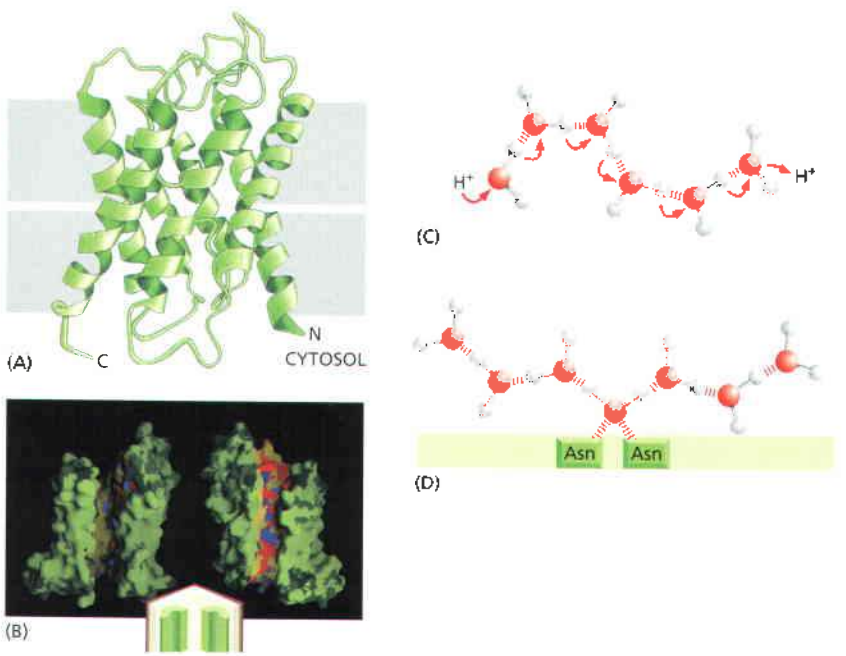


Figure 11–27 The structure of aquaporins. (A) A ribbon diagram of an aquaporin monomer. In the membrane, aquaporins form tetramers, with each monomer containing a pore in its center (not shown). (B) A space-filling model of an aquaporin monomer, which has been cut and opened like a book, so that the inside of the pore is visible. Hydrophilic amino acids lining the pores are colored *red* and *blue*, whereas hydrophobic amino acids lining the pore are colored *yellow*. The amino acids not involved in forming the pore are shown in *green*. Note that one face of the pore is lined with hydrophilic amino acids, which provide transient hydrogen bonds to water molecules; these bonds help line up the transiting water molecules in a single row and orient them as they traverse the membrane. By contrast, the other side of the pore is devoid of such amino acids, providing a *hydrophobic slide* that does not allow hydrogen bonds to form. (C and D) A model explaining why aquaporins are impermeable to H^+. (C) In water, H^+ diffuses extremely rapidly by being relayed from one water molecule to the next. (D) Two strategically placed asparagines in the center of each aquaporin pore help tether a central water molecule such that both valencies on its oxygen are occupied, thereby preventing an H^+ relay. (A and B, adapted from R.M. Stroud et al., *Curr. Opin. Struct. Biol.* 13:424–431, 2003. With

which diffuses through water extremely rapidly, using a molecular relay mechanism that requires the making and breaking of hydrogen bonds between adjacent water molecules (Figure 11–27C). Aquaporins contain two strategically placed asparagines, which bind to the oxygen atom of the central water molecule in the line of water molecules traversing the pore. Because both valences of this oxygen are unavailable for hydrogen bonding, the central water molecule cannot participate in an H^+ relay, and the pore is therefore impermeable to protons (Figure 11–27C and D).

Some bacterial water channels similar to aquaporins also conduct glycerol and small sugars, which interact with similarly positioned carbonyl oxygens lining the pore. Such transient contacts that solutes make with the pore walls ensure that the transport is highly specific, without significantly impeding the speed with which the solute passes. Each individual aquaporin channel passes about 10^9 water molecules per second.

The cells that make most sophisticated use of channels are neurons. Before discussing how they do so, we digress briefly to describe how a typical neuron is organized.

The Function of a Neuron Depends on Its Elongated Structure

The fundamental task of a **neuron**, or **nerve cell**, is to receive, conduct, and transmit signals. To perform these functions, neurons are often extremely elongated. A single neuron in a human, extending, for example, from the spinal cord to a muscle in the foot, may be as long as 1 meter. Every neuron consists of a cell body (containing the nucleus) with a number of thin processes radiating outward from it. Usually one long **axon** conducts signals away from the cell body toward distant targets, and several shorter branching **dendrites** extend from the cell body like antennae, providing an enlarged surface area to receive signals from the axons of other neurons (**Figure 11–28**), although the cell body itself also receives signals. A typical axon divides at its far end into many branches, passing on its message to many target cells simultaneously. Likewise, the extent of branching of the dendrites can be very great—in some cases sufficient to receive as many as 100,000 inputs on a single neuron.

Despite the varied significance of the signals carried by different classes of neurons, the form of the signal is always the same, consisting of changes in the electrical potential across the neuron's plasma membrane. The signal spreads because an electrical disturbance produced in one part of the cell spreads to other parts, although the disturbance becomes weaker with increasing distance from its source, unless the neuron expends energy to amplify it as it travels. Over short distances this attenuation is unimportant; in fact, many small neurons conduct their signals passively, without amplification. For long-distance communication, however, such passive spread is inadequate. Thus, larger neurons employ an active signaling mechanism, which is one of their most striking features. An electrical stimulus that exceeds a certain threshold strength triggers an explosion of electrical activity that propagates rapidly along the neuron's plasma

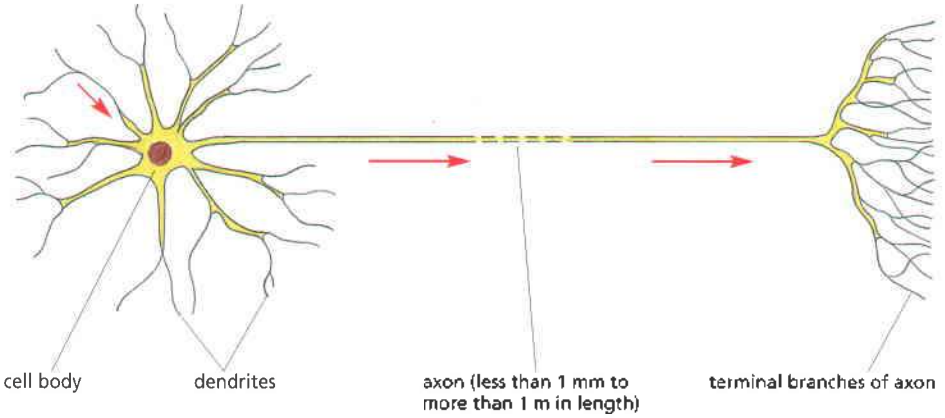


Figure 11–28 **A typical vertebrate neuron.** The arrows indicate the direction in which signals are conveyed. The single axon conducts signals away from the cell body, while the multiple dendrites (and the cell body) receive signals from the axons of other neurons. The nerve terminals end on the dendrites or cell body of other neurons or on other cell types, such as muscle or gland cells.

membrane and is sustained by automatic amplification all along the way. This traveling wave of electrical excitation, known as an **action potential**, or *nerve impulse*, can carry a message without attenuation from one end of a neuron to the other at speeds of 100 meters per second or more. Action potentials are the direct consequence of the properties of voltage-gated cation channels, as we now discuss.

Voltage-Gated Cation Channels Generate Action Potentials in Electrically Excitable Cells

The plasma membrane of all electrically excitable cells—not only neurons, but also muscle, endocrine, and egg cells—contains **voltage-gated cation channels**, which are responsible for generating the action potentials. An action potential is triggered by a *depolarization* of the plasma membrane—that is, by a shift in the membrane potential to a less negative value inside. (We shall see later how the action of a neurotransmitter causes depolarization.) In nerve and skeletal muscle cells, a stimulus that causes sufficient depolarization promptly opens the **voltage-gated Na^+ channels**, allowing a small amount of Na^+ to enter the cell down its electrochemical gradient. The influx of positive charge depolarizes the membrane further, thereby opening more Na^+ channels, which admit more Na^+ ions, causing still further depolarization. This self-amplification process (an example of *positive feedback*, discussed in Chapter 15), continues until, within a fraction of a millisecond, the electrical potential in the local region of membrane has shifted from its resting value of about –70 mV to almost as far as the Na^+ equilibrium potential of about +50 mV (see Panel 11–2, p. 670). At this point, when the net electrochemical driving force for the flow of Na^+ is almost zero, the cell would come to a new resting state, with all of its Na^+ channels permanently open, if the open conformation of the channel were stable. Two mechanisms that act in concert to save the cell from such a permanent electrical spasm: the Na^+ channels inactivate and voltage-gated K^+ channels open.

The Na^+ channels have an automatic inactivating mechanism, which causes the channels to reclose rapidly even though the membrane is still depolarized. The Na^+ channels remain in this *inactivated* state, unable to reopen, until after the membrane potential has returned to its initial negative value. The whole cycle from initial stimulus to the return to the original resting state takes a few milliseconds or less. The Na^+ channel can therefore exist in three distinct states—closed, open, and inactivated. **Figure 11–29** shows how the changes in state contribute to the rise and fall of the action potential.

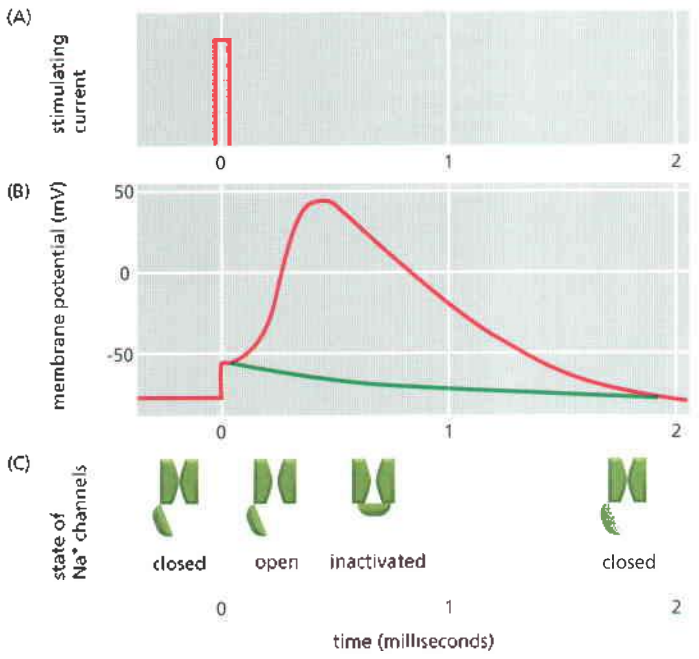
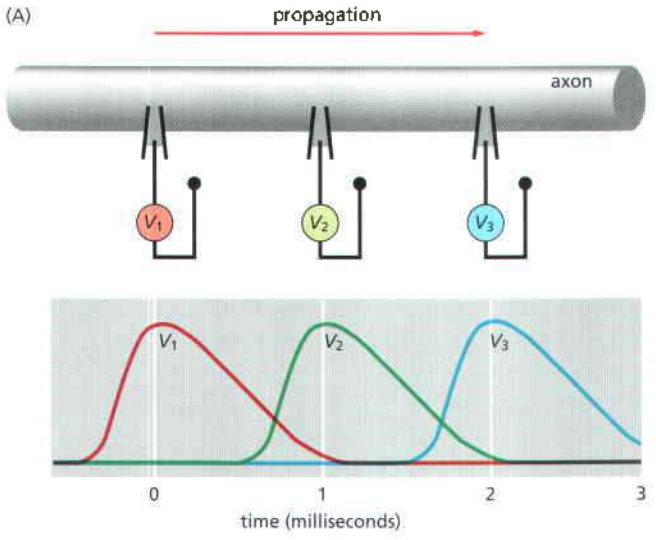


Figure 11–29 An action potential. <CGAG> (A) An action potential is triggered by a brief pulse of current, which (B) partially depolarizes the membrane, as shown in the plot of membrane potential versus time. The *green curve* shows how the membrane potential would have simply relaxed back to the resting value after the initial depolarizing stimulus if there had been no voltage-gated Na^+ channels in the membrane; this relatively slow return of the membrane potential to its initial value of –70 mV in the absence of open Na^+ channels occurs because of the efflux of K^+ through voltage-gated K^+ channels, which open in response to membrane depolarization and drive the membrane back toward the K^+ equilibrium potential. The *red curve* shows the course of the action potential that is caused by the opening and subsequent inactivation of voltage-gated Na^+ channels, whose state is shown in (C). The membrane cannot fire a second action potential until the Na^+ channels have returned to the closed conformation; until then, the membrane is refractory to stimulation.



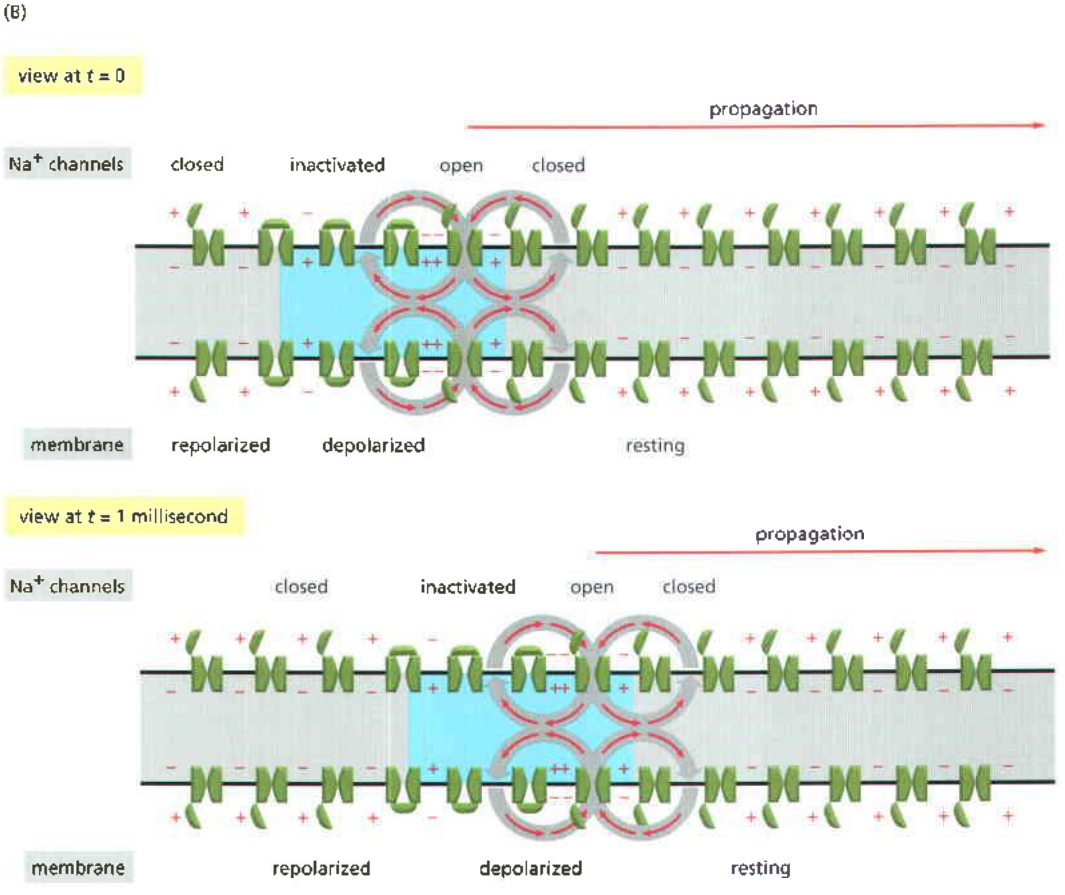


Figure 11–30 The propagation of an action potential along an axon. (A) The voltages that would be recorded from a set of intracellular electrodes placed at intervals along the axon. (B) The changes in the Na^+ channels and the current flows *(orange arrows)* that give rise to the traveling disturbance of the membrane potential. The region of the axon with a depolarized membrane is shaded in *blue*. Note that an action potential can only travel away from the site of depolarization, because Na^+-channel inactivation prevents the depolarization from spreading backward.

This description of an action potential applies only to a small patch of plasma membrane. The self-amplifying depolarization of the patch, however, is sufficient to depolarize neighboring regions of membrane, which then go through the same cycle. In this way, the action potential sweeps like a wave from the initial site of depolarization over the entire plasma membrane, as shown in Figure 11–30.

Voltage-gated K^+ channels provide a second mechanism in most nerve cells to help bring the activated plasma membrane more rapidly back toward its original negative potential, ready to transmit a second impulse. These channels open in response to membrane depolarization in much the same way that the Na^+ channels do, but with slightly slower kinetics; for this reason they are some-

times called *delayed K^+ channels*. Once the K^+ channels open, the efflux of K^+ rapidly overwhelms the transient influx of Na^+ and quickly drives the membrane back toward the K^+ equilibrium potential, even before the inactivation of the Na^+ channels is complete.

Like the Na^+ channel, the voltage-gated K^+ channels automatically inactivate. Studies of mutant voltage-gated K^+ channels show that the N-terminal 20 amino acids of the channel protein are required for rapid inactivation: altering this region changes the kinetics of channel inactivation, and removal of the region abolishes inactivation. Amazingly, in the latter case, exposing the cytoplasmic face of the plasma membrane to a small synthetic peptide corresponding to the missing N-terminus restores inactivation. These findings suggest that the N-terminus of each K^+ channel subunit acts like a tethered ball that occludes the cytoplasmic end of the pore soon after it opens, thereby inactivating the channel (**Figure 11–31**). A similar mechanism is thought to operate in the rapid inactivation of voltage-gated Na^+ channels (which we discuss later), although a different segment of the protein seems to be involved.

The electrochemical mechanism of the action potential was first established by a famous series of experiments carried out in the 1940s and 1950s. Because the techniques for studying electrical events in small cells had not yet been developed, the experiments exploited the giant neurons in the squid. Despite the many technical advances made since then, the logic of the original analysis continues to serve as a model for present-day work. **Panel 11–3** summarizes some of the key original experiments.

Myelination Increases the Speed and Efficiency of Action Potential Propagation in Nerve Cells

The axons of many vertebrate neurons are insulated by a **myelin sheath**, which greatly increases the rate at which an axon can conduct an action potential. The importance of myelination is dramatically demonstrated by the demyelinating disease *multiple sclerosis*, in which the immune system destroys myelin sheaths in some regions of the central nervous system; in the affected regions, the propagation of nerve impulses is greatly slowed, often with devastating neurological consequences.

Myelin is formed by specialized supporting cells, called **glial cells**. **Schwann cells** myelinate axons in peripheral nerves, and **oligodendrocytes** do so in the central nervous system. These glial cells wrap layer upon layer of their own plasma membrane in a tight spiral around the axon (**Figure 11–32A** and B), thereby insulating the axonal membrane so that little current can leak across it.

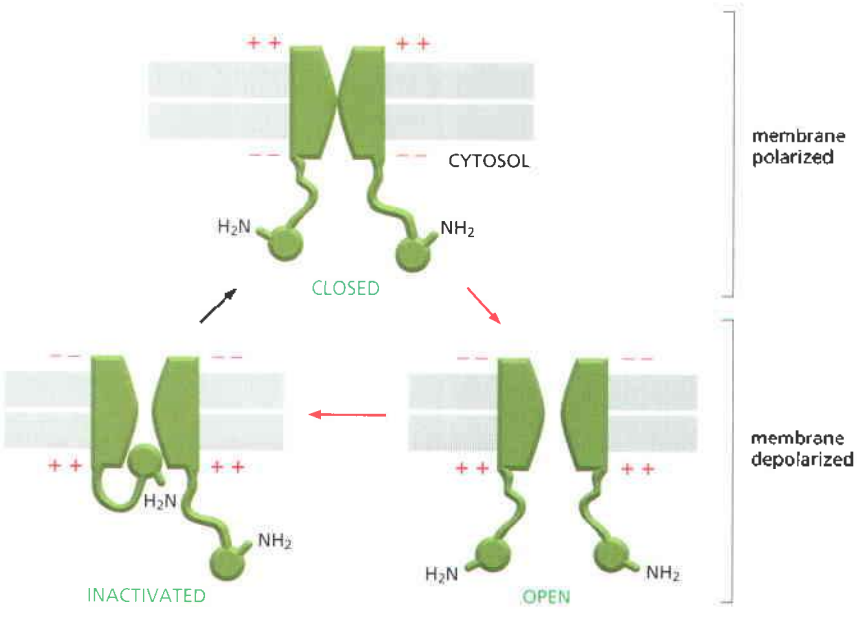
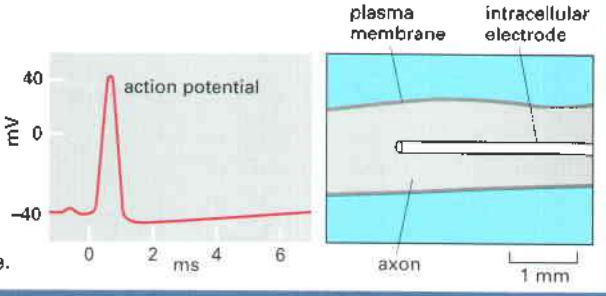


Figure 11–31 The "ball-and-chain" model of rapid inactivation of a voltage-gated K^+ channel. When the membrane is depolarized, the channel opens and begins to conduct ions. If the depolarization is maintained, the open channel adopts an inactive conformation, in which the pore is occluded by the N-terminal 20 amino acid "ball," which is linked to the channel proper by a segment of unfolded polypeptide chain that serves as the "chain." For simplicity, only two balls are shown; in fact, there are four, one from each subunit. A similar mechanism, using a different segment of the polypeptide chain, is thought to operate in Na^+ channel inactivation. Internal forces stabilize each state against small disturbances, but a sufficiently violent collision with other molecules can cause the channel to flip from one of these states to another. The state of lowest energy depends on the membrane potential because the different conformations have different charge distributions. When the membrane is at rest (highly polarized), the closed conformation has the lowest free energy and is therefore most stable; when the membrane is depolarized, the energy of the *open* conformation is lower, so the channel has a high probability of opening. But the free energy of the *inactivated* conformation is lower still; therefore, after a randomly variable period spent in the open state, the channel becomes inactivated. Thus, the open conformation corresponds to a metastable state that can exist only transiently. The *red arrows* indicate the sequence that follows a sudden depolarization; the *black arrow* indicates the return to the original conformation as the lowest energy state after the membrane is repolarized.

1. Action potentials are recorded with an intracellular electrode

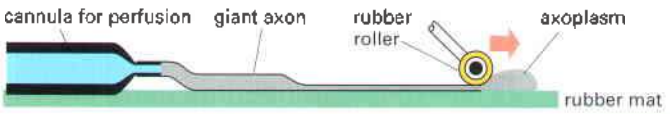
The squid giant axon is about 0.5–1 mm in diameter and several centimeters long. An electrode in the form of a glass capillary tube containing a conducting solution can be thrust down the axis of the axon so that its tip lies deep in the cytoplasm. With its help, one can measure the voltage difference between the inside and the outside of the axon—that is, the membrane potential—as an action potential sweeps past the electrode. The action potential is triggered by a brief electrical stimulus to one end of the axon. It does not matter which end, because the excitation can travel in either direction; and it does not matter how big the stimulus is, as long as it exceeds a certain threshold: the action potential is all or none.

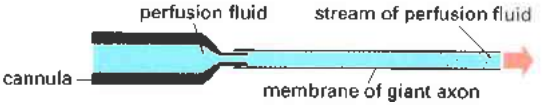


2. Action potentials depend only on the neuronal plasma membrane and on gradients of Na^+ and K^+ across it

The three most plentiful ions, both inside and outside the axon, are Na^+, K^+, and Cl^-. As in other cells, the Na^+-K^+ pump maintains a concentration gradient: the concentration of Na^+ is about 9 times lower inside the axon than outside, while the concentration of K^+ is about 20 times higher inside than outside. Which ions are important for the action potential?

The squid giant axon is so large and robust that it is possible to extrude the gel-like cytoplasm from it, like toothpaste from a tube, and then to perfuse it internally with pure artificial solutions of Na^+, K^+, and Cl^- or SO_4^{2-}. Remarkably, if (and only if) the concentrations of Na^+ and K^+ inside and outside approximate those found naturally, the axon will still propagate action potentials of the normal form. The important part of the cell for electrical signaling, therefore, must be the plasma membrane; the important ions are Na^+ and K^+; and a sufficient source of free energy to power the action potential must be provided by the concentration gradients of these ions across the membrane, because all other sources of metabolic energy have presumably been removed by the perfusion.

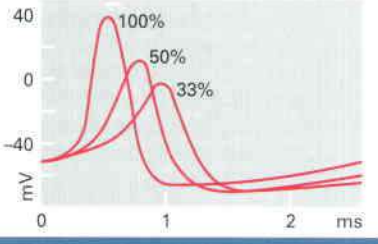




3. At rest, the membrane is chiefly permeable to K^+; during the action potential, it becomes transiently permeable to Na^+

At rest the membrane potential is close to the equilibrium potential for K^+. When the external concentration of K^+ is changed, the resting potential changes roughly in accordance with the Nernst equation for K^+ (see Panel 11–2). At rest, therefore, the membrane is chiefly permeable to K^+: K^+ leak channels provide the main ion pathway through the membrane.

If the external concentration of Na^+ is varied, there is no effect on the resting potential. However, the height of the peak of the action potential varies roughly in accordance with the Nernst equation for Na^+. During the action potential, therefore, the membrane appears to be chiefly permeable to Na^+: Na^+ channels have opened. In the aftermath of the action potential, the membrane potential reverts to a negative value that depends on the external concentration of K^+ and is even closer to the K^+ equilibrium potential than the resting potential is: the membrane has lost most of its permeability to Na^+ and has become even more permeable to K^+ than before—that is, Na^+ channels have closed, and additional K^+ channels have opened.

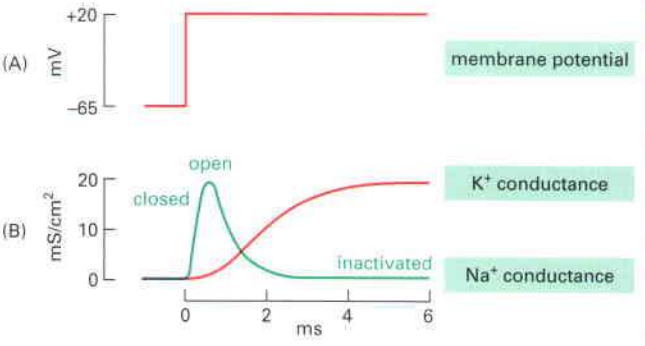


The form of the action potential when the external medium contains 100%, 50%, or 33% of the normal concentration of Na^+.

4. Voltage clamping reveals how the membrane potential controls opening and closing of ion channels

The membrane potential can be held constant ("voltage clamped") throughout the axon by passing a suitable current through a bare metal wire inserted along the axis of the axon while monitoring the membrane potential with another intracellular electrode. When the membrane is abruptly shifted from the resting potential and held in a depolarized state (A), Na^+ channels rapidly open until the Na^+ permeability of the membrane is much greater than the K^+ permeability; they then close again spontaneously, even though the membrane potential is clamped and unchanging. K^+ channels also open but with a delay, so that the K^+ permeability increases as the Na^+ permeability falls (B). If the experiment is now very promptly repeated, by returning the membrane briefly to the resting potential and then quickly depolarizing it again, the response is different: prolonged depolarization has caused the Na^+ channels to enter an inactivated state, so that the second depolarization fails to cause a rise and fall similar to the first. Recovery from this state requires a relatively long time—about 10 milliseconds—spent at the repolarized (resting) membrane potential.

In a normal unclamped axon, an inrush of Na^+ through the opened Na^+ channels produces the spike of the action potential; inactivation of Na^+ channels and opening of K^+ channels bring the membrane rapidly back down to the resting potential.



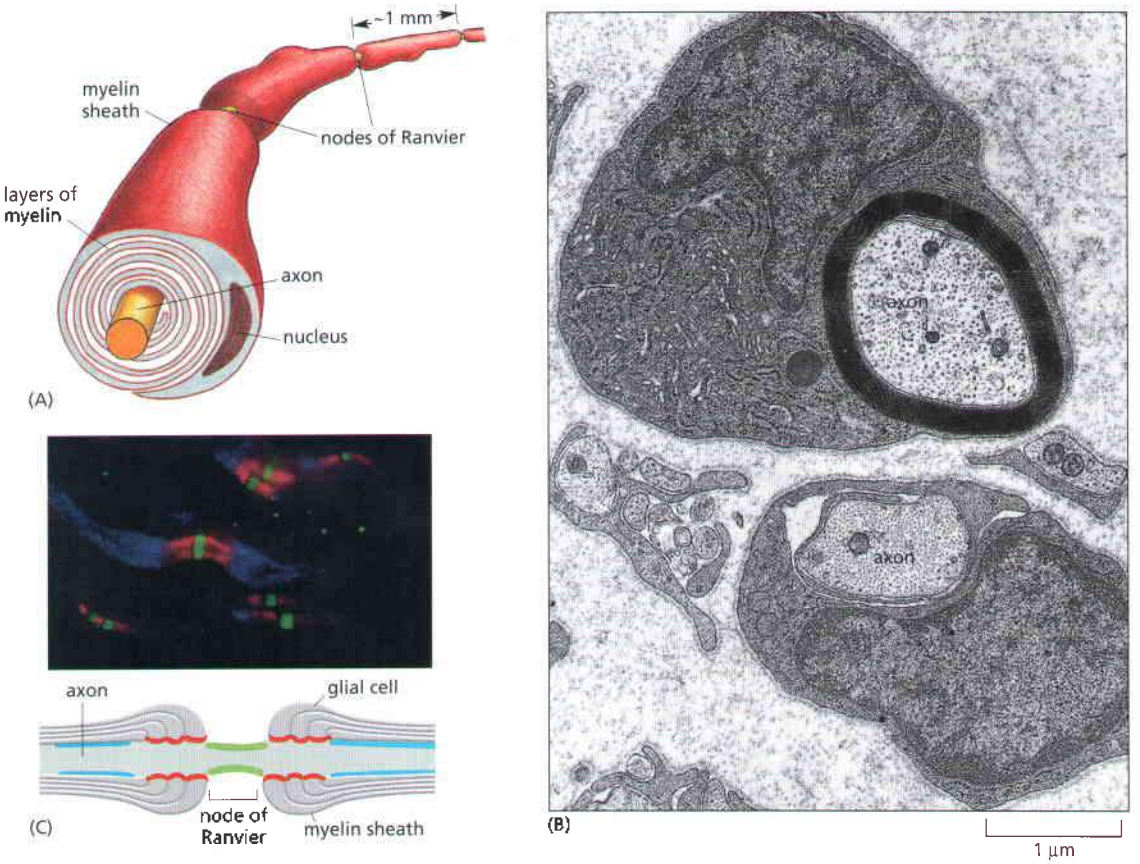


Figure 11–32 Myelination. (A) A myelinated axon from a peripheral nerve. Each Schwann cell wraps its plasma membrane concentrically around the axon to form a segment of myelin sheath about 1 mm long. For clarity, the membrane layers of the myelin in this drawing are shown less compacted than they are in reality (see part B). (B) An electron micrograph of a section from a nerve in the leg of a young rat. Two Schwann cells can be seen: one near the bottom is just beginning to myelinate its axon; the one above it has formed an almost mature myelin sheath. (C) Fluorescence micrograph and diagram of individual myelinated axons teased apart in a nerve. Three different proteins are detected by staining with antibodies. Voltage-gated Na^+ channels (stained in *green*) are concentrated in the axonal membrane at the nodes of Ranvier. An extracellular protein (called *Caspr*, stained in *red*) marks the end of each myelin sheath. Caspr assembles at the junctions where the glial cell plasma membrane tightly abuts the axon to provide the electrical seal. Voltage-gated K^+ channels (stained in *blue*) localize to regions in the axon plasma membrane that are close to the nodes. (B, from Cedric S. Raine, in Myelin [P. Morell, ed.]. New York: Plenum, 1976; C, from M.N. Rasband and P. Shrager, *J. Physiol.* 525:63–73, 2000. With permission from Blackwell Publishing.)

The myelin sheath is interrupted at regularly spaced *nodes of Ranvier*, where almost all the Na^+ channels in the axon are concentrated (see Figure 11–32C). Because the ensheathed portions of the axonal membrane have excellent cable properties (in other words, they behave electrically much like well-designed underwater telegraph cables), a depolarization of the membrane at one node almost immediately spreads passively to the next node. Thus, an action potential propagates along a myelinated axon by jumping from node to node, a process called *saltatory conduction*. This type of conduction has two main advantages: action potentials travel faster, and metabolic energy is conserved because the active excitation is confined to the small regions of axonal plasma membrane at nodes of Ranvier.

Patch-Clamp Recording Indicates That Individual Gated Channels Open in an All-or-Nothing Fashion

Neuron and skeletal muscle cell plasma membranes contain many thousands of voltage-gated Na^+ channels, and the current crossing the membrane is the sum of the currents flowing through all of these. An intracellular microelectrode can

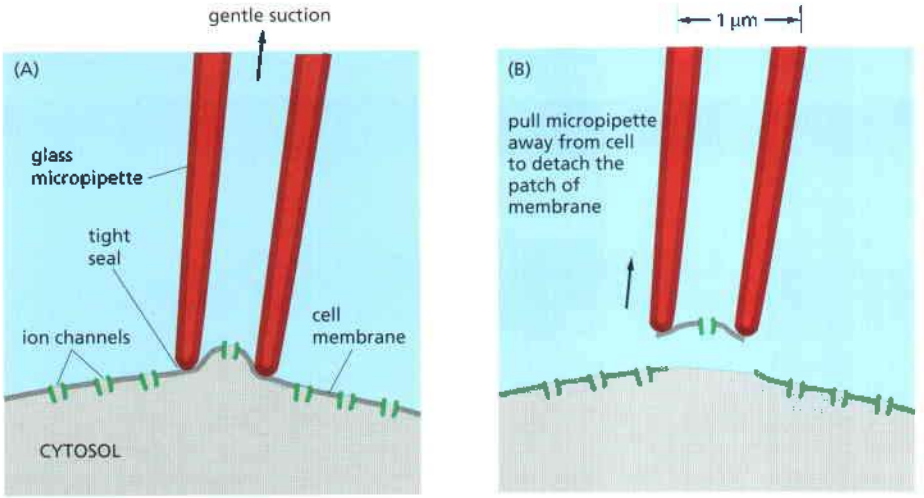


Figure 11–33 The technique of patch-clamp recording. Because of the extremely tight seal between the micropipette and the membrane, current can enter or leave the micropipette only by passing through the channels in the *patch* of membrane covering its tip. The term *clamp* is used because an electronic device is employed to maintain, or "clamp," the membrane potential at a set value while recording the ionic current through individual channels. The current through these channels can be recorded with the patch still attached to the rest of the cell, as in (A), or detached, as in (B). The advantage of the detached patch is that it is easy to alter the composition of the solution on either side of the membrane to test the effect of various solutes on channel behavior. A detached patch can also be produced with the opposite orientation, so that the cytoplasmic surface of the membrane faces the inside of the pipette.

record this aggregate current, as shown in Figure 11–30. Remarkably, however, it is also possible to record current flowing through individual channels. **Patch-clamp recording**, developed in the 1970s and 80s has revolutionized the study of ion channels, making it possible to examine transport through a single molecule of channel protein in a small patch of membrane covering the mouth of a micropipette (**Figure 11–33**). With this simple but powerful technique, one can study the detailed properties of ion channels in all sorts of cell types. This work has led to the discovery that even cells that are not electrically excitable usually have a variety of gated ion channels in their plasma membrane. Many of these cells, such as yeasts, are too small to be investigated by the traditional electrophysiologist's method of impalement with an intracellular microelectrode.

Patch-clamp recording indicates that individual voltage-gated Na^+ channels open in an all-or-nothing fashion. A channel opens and closes at random, but when open, the channel always has the same large conductance, allowing more than 1000 ions to pass per millisecond. Therefore, the aggregate current crossing the membrane of an entire cell does not indicate the *degree* to which a typical individual channel is open but rather the *total number* of channels in its membrane that are open at any one time (**Figure 11–34**).

Some simple physical principles allow us to understand voltage-gating. The interior of the resting neuron or muscle cell is at an electrical potential about 50–100 mV more negative than the external medium. Although this potential difference seems small, it exists across a plasma membrane only about 5 nm thick, so that the resulting voltage gradient is about 100,000 V/cm. Proteins in the

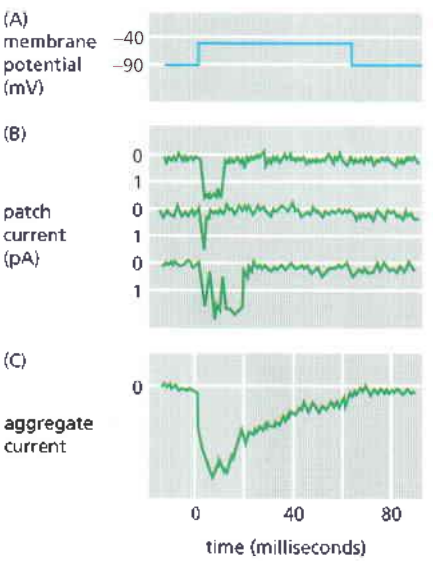


Figure 11–34 Patch-clamp measurements for a single voltage-gated Na^+ channel. A tiny patch of plasma membrane was detached from an embryonic rat muscle cell, as in Figure 11–33. (A) The membrane was depolarized by an abrupt shift of potential. (B) Three current records from three experiments performed on the same patch of membrane. Each major current step in (B) represents the opening and closing of a single channel. A comparison of the three records shows that, whereas the durations of channel opening and closing vary greatly, the rate at which current flows through an open channel is practically constant. The minor fluctuations in the current records arise largely from electrical noise in the recording apparatus. Current is measured in picoamperes (pA). By convention, the electrical potential on the outside of the cell is defined as zero. (C) The sum of the currents measured in 144 repetitions of the same experiment. This aggregate current is equivalent to the usual Na^+ current that would be observed flowing through a relatively large region of membrane containing 144 channels. A comparison of (B) and (C) reveals that the time course of the aggregate current reflects the probability that any individual channel will be in the open state; this probability decreases with time as the channels in the depolarized membrane adopt their inactivated conformation. (Data from J. Patlak and R. Horn, *J. Gen. Physiol.* 79:333–351, 1982. With permission from The Rockefeller University Press.)

membrane are thus subjected to a very large electrical field that can profoundly affect their conformation. These proteins, like all others, have many charged groups, as well as polarized bonds between their various atoms. The electrical field therefore exerts forces on the molecular structure. For many membrane proteins the effects of changes in the membrane electrical field are probably insignificant, but voltage-gated ion channels can adopt alternative conformations whose stabilities depend on the strength of the field. Voltage-gated Na^+, K^+, and Ca^{2+} channels, for example, have characteristic positively charged amino acids in one of their transmembrane segments that respond to depolarization by moving outward, triggering conformational changes that open the channel. Each conformation can "flip" to another conformation if given a sufficient jolt by the random thermal movements of the surroundings, and it is the relative stability of the closed, open, and inactivated conformations against flipping that is altered by changes in the membrane potential (see legend to Figure 11–31).

Voltage-Gated Cation Channels Are Evolutionarily and Structurally Related

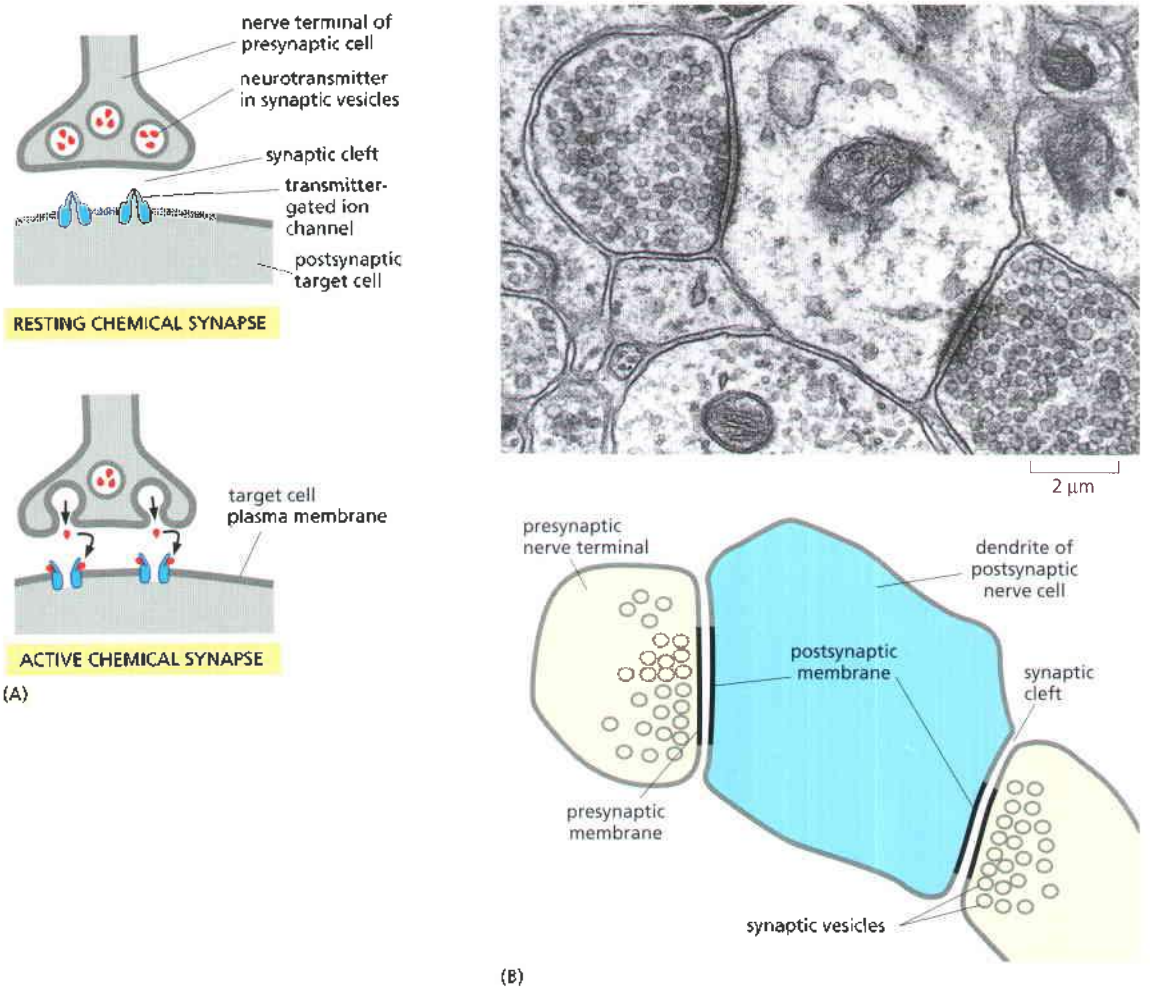
Na^+ channels are not the only kind of voltage-gated cation channel that can generate an action potential. The action potentials in some muscle, egg, and endocrine cells, for example, depend on *voltage-gated Ca^{2+} channels* rather than on Na^+ channels.

There is a surprising amount of structural and functional diversity within each of these three classes, generated both by multiple genes and by the alternative splicing of RNA transcripts produced from the same gene. Nonetheless, the amino acid sequences of the known voltage-gated Na^+, K^+, and Ca^{2+} channels show striking similarities, demonstrating that they all belong to a large superfamily of evolutionarily and structurally related proteins and share many of the design principles. Whereas the single-celled yeast *S. cerevisiae* contains a single gene that codes for a voltage-gated K^+ channel, the genome of the worm *C. elegans* contains 68 genes that encode different but related K^+ channels. This complexity indicates that even a simple nervous system made up of only 302 neurons uses a large number of different ion channels to compute its responses.

Humans who inherit mutant genes encoding ion channel proteins can suffer from a variety of nerve, muscle, brain, or heart diseases, depending in which cells the channel encoded by the mutant gene normally functions. Mutations in genes that encode voltage-gated Na^+ channels in skeletal muscle cells, for example, can cause *myotonia*, a condition in which there is a delay in muscle relaxation after voluntary contraction, causing painful muscle spasms. In some cases, this occurs because the abnormal channels fail to inactivate normally; as a result, Na^+ entry persists after an action potential finishes and repeatedly reinitiates membrane depolarization and muscle contraction. Similarly, mutations that affect Na^+ or K^+ channels in the brain can cause *epilepsy*, in which excessive synchronized firing of large groups of neurons cause epileptic seizures (convulsions, or fits).

Transmitter-Gated Ion Channels Convert Chemical Signals into Electrical Ones at Chemical Synapses

Neuronal signals are transmitted from cell to cell at specialized sites of contact known as **synapses**. The usual mechanism of transmission is indirect. The cells are electrically isolated from one another, the *presynaptic cell* being separated from the *postsynaptic cell* by a narrow *synaptic cleft*. A change of electrical potential in the presynaptic cell triggers it to release small signal molecules known as **neurotransmitters**, which are stored in membrane-enclosed synaptic vesicles and released by exocytosis. The neurotransmitter diffuses rapidly across the synaptic cleft and provokes an electrical change in the postsynaptic cell by binding to *transmitter-gated ion channels* (**Figure 11–35**) and opening them. After the neurotransmitter has been secreted, it is rapidly removed: it is either



destroyed by specific enzymes in the synaptic cleft or taken up by the nerve terminal that released it or by surrounding glial cells. Reuptake is mediated by a variety of Na^+-dependent neurotransmitter transporters; in this way, neurotransmitters are recycled, allowing cells to keep up with high rates of release. Rapid removal ensures both spatial and temporal precision of signaling at a synapse. It decreases the chances that the neurotransmitter will influence neighboring cells, and it clears the synaptic cleft before the next pulse of neurotransmitter is released, so that the timing of repeated, rapid signaling events can be accurately communicated to the postsynaptic cell. As we shall see, signaling via such *chemical synapses* is far more versatile and adaptable than direct electrical coupling via gap junctions at *electrical synapses* (discussed in Chapter 19), which are also used by neurons but to a much smaller extent.

Transmitter-gated ion channels are specialized for rapidly converting extracellular chemical signals into electrical signals at chemical synapses. The channels are concentrated in the plasma membrane of the postsynaptic cell in the region of the synapse and open transiently in response to the binding of neurotransmitter molecules, thereby producing a brief permeability change in the membrane (see Figure 11–35A). Unlike the voltage-gated channels responsible for action potentials, transmitter-gated channels are relatively insensitive to the membrane potential and therefore cannot by themselves produce a self-amplifying excitation. Instead, they produce local permeability changes, and hence changes of membrane potential, that are graded according to the amount of neurotransmitter released at the synapse and how long it persists there. An action potential can be triggered from this site only if the local membrane potential depolarizes enough to open a sufficient number of nearby voltage-gated cation channels that are present in the same target cell membrane.

Figure 11–35 A chemical synapse. <CTGA> (A) When an action potential reaches the nerve terminal in a presynaptic cell, it stimulates the terminal to release its neurotransmitter. The neurotransmitter molecules are contained in synaptic vesicles and are released to the cell exterior when the vesicles fuse with the plasma membrane of the nerve terminal. The released neurotransmitter binds to and opens the transmitter-gated ion channels concentrated in the plasma membrane of the postsynaptic target cell at the synapse. The resulting ion flows alter the membrane potential of the target cell, thereby transmitting a signal from the excited nerve. (B) A thin section electron micrograph of two nerve terminal synapses on a dendrite of a postsynaptic cell. (B, courtesy of Cedric Raine.)

Chemical Synapses Can Be Excitatory or Inhibitory

Transmitter-gated ion channels differ from one another in several important ways. First, as receptors, they have a highly selective binding site for the neurotransmitter that is released from the presynaptic nerve terminal. Second, as channels, they are selective in the type of ions that they let pass across the plasma membrane; this determines the nature of the postsynaptic response. **Excitatory neurotransmitters** open cation channels, causing an influx of Na^+ that depolarizes the postsynaptic membrane toward the threshold potential for firing an action potential. **Inhibitory neurotransmitters**, by contrast, open either Cl^- channels or K^+ channels, and this suppresses firing by making it harder for excitatory influences to depolarize the postsynaptic membrane. Many transmitters can be either excitatory or inhibitory, depending on where they are released, what receptors they bind to, and the ionic conditions that they encounter. *Acetylcholine*, for example, can either excite or inhibit, depending on the type of acetylcholine receptors it binds to. Usually, however, acetylcholine, *glutamate*, and *serotonin* are used as excitatory transmitters, and *γ-aminobutyric acid (GABA)* and *glycine* are used as inhibitory transmitters. Glutamate, for instance, mediates most of the excitatory signaling in the vertebrate brain.

We have already discussed how the opening of cation channels depolarizes a membrane. We can understand the effect of opening Cl^- channels as follows. The concentration of Cl^- is much higher outside the cell than inside (see Table 11–1, p. 652), but the membrane potential opposes its influx. In fact, for many neurons, the equilibrium potential for Cl^- is close to the resting potential—or even more negative. For this reason, opening Cl^- channels tends to buffer the membrane potential; as the membrane starts to depolarize, more negatively charged Cl^- ions enter the cell and counteract the depolarizatrion. Thus, the opening of Cl^- channels makes it more difficult to depolarize the membrane and hence to excite the cell. The opening of K^+ channels has a similar effect. The effects of toxins that block their action demonstrate the importance of inhibitory neurotransmitters: strychnine, for example, by binding to glycine receptors and blocking the inhibitory action of glycine, causes muscle spasms, convulsions, and death.

However, not all chemical signaling in the nervous system operates through ligand-gated ion channels. Many of the signaling molecules that are secreted by nerve terminals, including a large variety of neuropeptides, bind to receptors that regulate ion channels only indirectly. We discuss these so-called *G-protein-coupled receptors* and *enzyme-coupled receptors* in detail in Chapter 15. Whereas signaling mediated by excitatory and inhibitory neurotransmitters binding to transmitter-gated ion channels is generally immediate, simple, and brief, signaling mediated by ligands binding to G-protein-coupled receptors and enzyme-coupled receptors tends to be far slower and more complex, and longer lasting in its consequences.

The Acetylcholine Receptors at the Neuromuscular Junction Are Transmitter-Gated Cation Channels

The best-studied example of a transmitter-gated ion channel is the *acetylcholine receptor* of skeletal muscle cells. This channel is opened transiently by acetylcholine released from the nerve terminal at a **neuromuscular junction**—the specialized chemical synapse between a motor neuron and a skeletal muscle cell (**Figure 11–36**). This synapse has been intensively investigated because it is readily accessible to electrophysiological study, unlike most of the synapses in the central nervous system.

The **acetylcholine receptor** has a special place in the history of ion channels. It was the first ion channel to be purified, the first to have its complete amino acid sequence determined, the first to be functionally reconstituted in synthetic lipid bilayers, and the first for which the electrical signal of a single open channel was recorded. Its gene was also the first ion channel gene to be cloned and sequenced, and its three-dimensional structure has been determined, albeit at

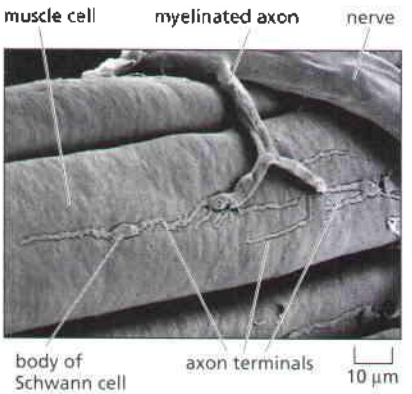


Figure 11–36 A low-magnification scanning electron micrograph of a neuromuscular junction in a frog. The termination of a single axon on a skeletal muscle cell is shown. (From J. Desaki and Y. Uehara, *J. Neurocytol.* 10:101–110, 1981. With permission from Kluwer Academic Publishers.)

only moderate resolution. There were at least two reasons for the rapid progress in purifying and characterizing this receptor. First, an unusually rich source of the acetylcholine receptors exists in the electric organs of electric fish and rays (these organs are modified muscles designed to deliver a large electric shock to prey). Second, certain neurotoxins (such as *α-bungarotoxin*) in the venom of certain snakes bind with high affinity ($K_a = 10^9$ liters/mole) and specificity to the receptor and can therefore be used to purify it by affinity chromatography. Fluorescent or radiolabeled α-bungarotoxin can also be used to localize and count acetylcholine receptors. In this way, researchers have shown that the receptors are densely packed in the muscle cell plasma membrane at a neuromuscular junction (about 20,000 such receptors per μm^2), with relatively few receptors elsewhere in the same membrane.

The acetylcholine receptor of skeletal muscle is composed of five transmembrane polypeptides, two of one kind and three others, encoded by four separate genes. The four genes are strikingly similar in sequence, implying that they evolved from a single ancestral gene. The two identical polypeptides in the pentamer each contribute to one of two binding sites for acetylcholine that are nestled between adjoining subunits. When two acetylcholine molecules bind to the pentameric complex, they induce a conformational change: the helices that line the pore rotate to disrupt a ring of hydrophobic amino acids that blocks ion flow in the closed state. With ligand bound, the channel still flickers between open and closed states, but now it has a 90% probability of being open. This state continues until hydrolysis by a specific enzyme *(acetylcholinesterase)* located at the neuromuscular junction lowers the concentration of acetylcholine sufficiently. Once freed of its bound neurotransmitter, the acetylcholine receptor reverts to its initial resting state. If the presence of acetylcholine persists for a prolonged time as a result of excessive nerve stimulation, the channel inactivates (**Figure 11–37**).

The general shape of the acetylcholine receptor and the likely arrangement of its subunits have been determined by electron microscopy (**Figure 11–38**). The five subunits are arranged in a ring, forming a water-filled transmembrane channel that consists of a narrow pore through the lipid bilayer, which widens into vestibules at both ends. Clusters of negatively charged amino acids at either end of the pore help to exclude negative ions and encourage any positive ion of diameter less than 0.65 nm to pass through. The normal traffic consists chiefly of Na^+ and K^+, together with some Ca^{2+}. Thus, unlike voltage-gated cation channels, such as the K^+ channel discussed earlier, there is little selectivity among cations, and the relative contributions of the different cations to the current through the channel depend chiefly on their concentrations and on the electrochemical driving forces. When the muscle cell membrane is at its resting potential, the net driving force for K^+ is near zero, since the voltage gradient nearly balances the K^+ concentration gradient across the membrane (see Panel 11–2, p. 670). For Na^+, in contrast, the voltage gradient and the concentration gradient both act in the same direction to drive the ion into the cell. (The same is true for Ca^{2+}, but the extracellular concentration of Ca^{2+} is so much lower than that of Na^+ that Ca^{2+} makes only a small contribution to the total inward current.)

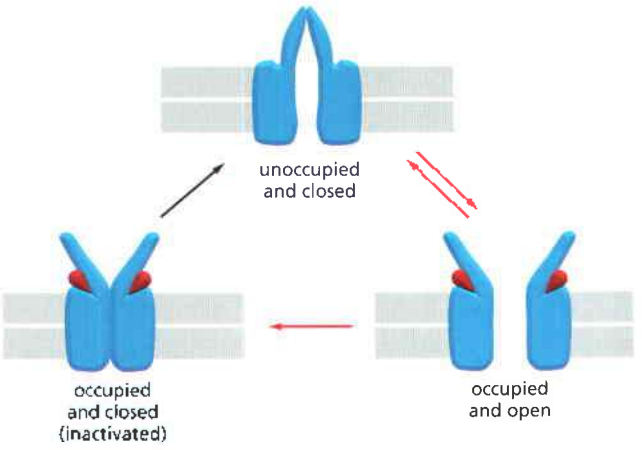


Figure 11–37 Three conformations of the acetylcholine receptor. The binding of two acetylcholine molecules opens this transmitter-gated ion channel. It then maintains a high probability of being open until the acetylcholine has been hydrolyzed. In the persistent presence of acetylcholine, however, the channel inactivates (desensitizes). Normally, the acetylcholine is rapidly hydrolyzed and the channel closes within about 1 millisecond, well before significant desensitization occurs. Desensitization would occur after about 20 milliseconds in the continued presence of acetylcholine.

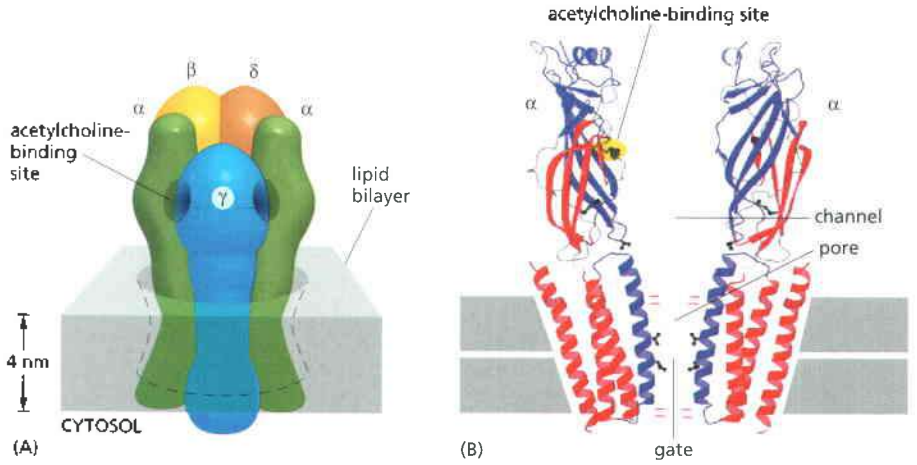


Figure 11–38 A model for the structure of the acetylcholine receptor. (A) Five homologous subunits (α, α, β, γ, δ) combine to form a transmembrane aqueous pore. The pore is lined by a ring of five transmembrane α helices, one contributed by each subunit. In its closed conformation, the pore is thought to be occluded by the hydrophobic side chains of five leucines, one from each α helix, which form a gate near the middle of the lipid bilayer. The negatively charged side chains at either end of the pore ensure that only positively charged ions pass through the channel. (B) Both of the α subunits contribute to an acetylcholine-binding site nestled between adjoining subunits; when acetylcholine binds to both sites, the channel undergoes a conformational change that opens the gate, possibly by rotating the helices containing the occluding leucines to move outward. In the structural drawing *(right)*, the parts of the channel that move in response to AChR binding to open the pore are colored in *blue*. (Adapted from N. Unwin, *Cell* 72[Suppl.]:31–41, 1993. With permission from Elsevier.)

Therefore, the opening of the acetylcholine receptor channels leads to a large net influx of Na^+ (a peak rate of about 30,000 ions per channel each millisecond). This influx causes a membrane depolarization that signals the muscle to contract, as discussed below.

Transmitter-Gated Ion Channels Are Major Targets for Psychoactive Drugs

The ion channels that open directly in response to the neurotransmitters acetylcholine, serotonin, GABA, and glycine contain subunits that are structurally similar and probably form transmembrane pores in the same way, even though they have distinct neurotransmitter-binding specificities and ion selectivities. These channels are all built from homologous polypeptide subunits, which probably assemble as a pentamer resembling the acetylcholine receptor. Glutamate-gated ion channels are constructed from a distinct family of subunits and are thought to form tetramers resembling the K^+ channels discussed earlier.

For each class of transmitter-gated ion channel, there are alternative forms of each type of subunit, either encoded by distinct genes or generated by alternative RNA splicing of the same gene product. The subunits assemble in different combinations to form an extremely diverse set of distinct channel subtypes, with different ligand affinities, different channel conductances, different rates of opening and closing, and different sensitivities to drugs and toxins. Vertebrate neurons, for example, have acetylcholine-gated ion channels that differ from those of muscle cells in that they are usually formed from two subunits of one type and three of another; but there are at least nine genes coding for different versions of the first type of subunit and at least three coding for different versions of the second, with further diversity due to alternative RNA splicing. Subsets of acetylcholine-sensitive neurons performing different functions in the brain express different combinations of these subunits. This, in principle, and already to some extent in practice, makes it possible to design drugs targeted against narrowly defined groups of neurons or synapses, thereby specifically influencing particular brain functions.

Indeed, transmitter-gated ion channels have for a long time been important targets for drugs. A surgeon, for example, can relax muscles for the duration of an operation by blocking the acetylcholine receptors on skeletal muscle cells with *curare*, a drug from a plant that was originally used by South American Indians to make poison arrows. Most drugs used to treat insomnia, anxiety, depression, and schizophrenia exert their effects at chemical synapses, and many of these act by binding to transmitter-gated channels. Both barbiturates and tranquilizers, such as Valium and Librium, for example, bind to GABA receptors, potentiating the inhibitory action of GABA by allowing lower concentrations of this neurotransmitter to open Cl^- channels. The new molecular biology of ion

channels, by revealing both their diversity and the details of their structure, holds out the hope of designing a new generation of psychoactive drugs that will act still more selectively to alleviate the miseries of mental illness.

In addition to ion channels, many other components of the synaptic signaling machinery are potential targets for psychoactive drugs. As mentioned earlier, after release into the synaptic cleft, many neurotransmitters are cleared by reuptake mechanisms mediated by Na^+-driven transporters. The inhibition of such a transporter prolongs the effect of the transmitter and thereby strengthens synaptic transmission. Many antidepressant drugs, including Prozac, for example, inhibit the uptake of serotonin; others inhibit the uptake of both serotonin and norepinephrine.

Ion channels are the basic molecular components from which neuronal devices for signaling and computation are built. To provide a glimpse of how sophisticated the functions of these devices can be, we consider several examples that demonstrate how groups of ion channels work together in synaptic communication between electrically excitable cells.

Neuromuscular Transmission Involves the Sequential Activation of Five Different Sets of Ion Channels

The following process, in which a nerve impulse stimulates a muscle cell to contract, illustrates the importance of ion channels to electrically excitable cells. This apparently simple response requires the sequential activation of at least five different sets of ion channels, all within a few milliseconds (**Figure 11–39**).

1. The process is initiated when the nerve impulse reaches the nerve terminal and depolarizes the plasma membrane of the terminal. The depolarization transiently opens voltage-gated Ca^{2+} channels in this membrane. As the Ca^{2+} concentration outside cells is more than 1000 times greater than the free Ca^{2+} concentration inside, Ca^{2+} flows into the nerve terminal. The increase in Ca^{2+} concentration in the cytosol of the nerve terminal triggers the local release of acetylcholine into the synaptic cleft.
2. The released acetylcholine binds to acetylcholine receptors in the muscle cell plasma membrane, transiently opening the cation channels associated with them. The resulting influx of Na^+ causes a local membrane depolarization.
3. The local depolarization of the muscle cell plasma membrane opens voltage-gated Na^+ channels in this membrane, allowing more Na^+ to enter, which further depolarizes the membrane. This, in turn, opens neighboring voltage-gated Na^+ channels and results in a self-propagating depolarization (an action potential) that spreads to involve the entire plasma membrane (see Figure 11–30).

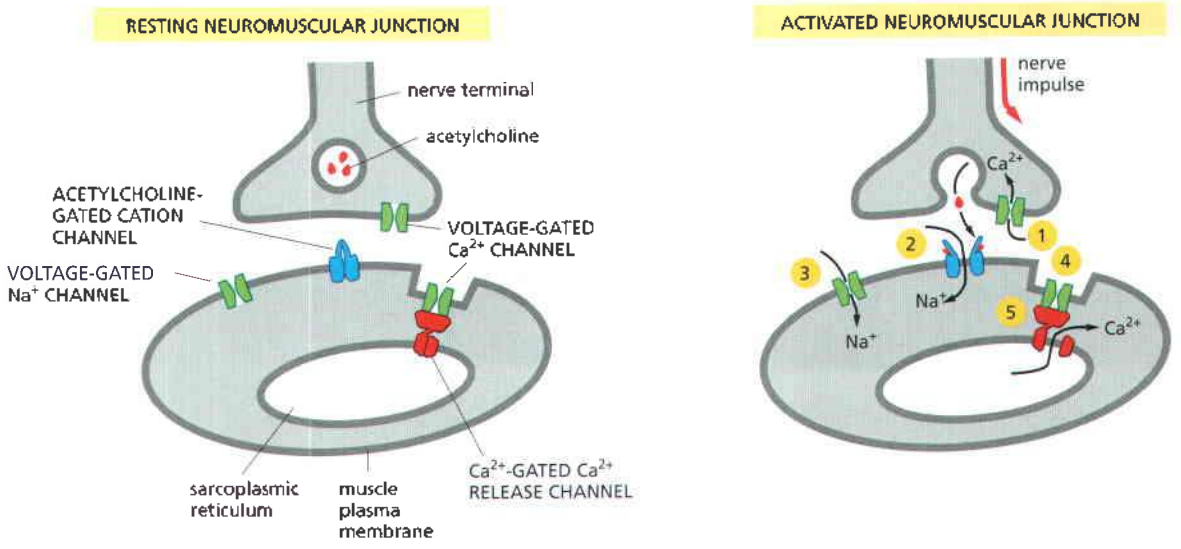


Figure 11–39 The system of ion channels at a neuromuscular junction. These gated ion channels are essential for the stimulation of muscle contraction by a nerve impulse. The various channels are numbered in the sequence in which they are activated, as described in the text.

4. The generalized depolarization of the muscle cell plasma membrane activates voltage-gated Ca^{2+} channels in specialized regions (the transverse [T] tubules—discussed in Chapter 16) of this membrane.
5. This, in turn, causes Ca^{2+}-gated *Ca^{2+} release channels* in an adjacent region of the sarcoplasmic reticulum (SR) membrane to open transiently and release the Ca^{2+} stored in the SR into the cytosol. The T-tubule and SR membranes are closely apposed with the two types of channels joined together in a specialized structure (see Figure 16–77). It is the sudden increase in the cytosolic Ca^{2+} concentration that causes the myofibrils in the muscle cell to contract.

Whereas the activation of muscle contraction by a motor neuron is complex, an even more sophisticated interplay of ion channels is required for a neuron to integrate a large number of input signals at synapses and compute an appropriate output, as we now discuss.

Single Neurons Are Complex Computation Devices

In the central nervous system, a single neuron can receive inputs from thousands of other neurons, and can in turn form synapses with many thousands of other cells. Several thousand nerve terminals, for example, make synapses on an average motor neuron in the spinal cord; its cell body and dendrites are almost completely covered with them (**Figure 11–40**). Some of these synapses transmit signals from the brain or spinal cord; others bring sensory information from muscles or from the skin. The motor neuron must combine the information received from all these sources and react either by firing action potentials along its axon or by remaining quiet.

Of the many synapses on a neuron, some tend to excite it, others to inhibit it. Neurotransmitter released at an excitatory synapse causes a small depolarization in the postsynaptic membrane called an *excitatory postsynaptic potential (excitatory PSP)*, while neurotransmitter released at an inhibitory synapse generally causes a small hyperpolarization called an *inhibitory PSP*. The membrane of the dendrites and cell body of most neurons contains a relatively low density of voltage-gated Na^+ channels, and an individual excitatory PSP is generally too small to

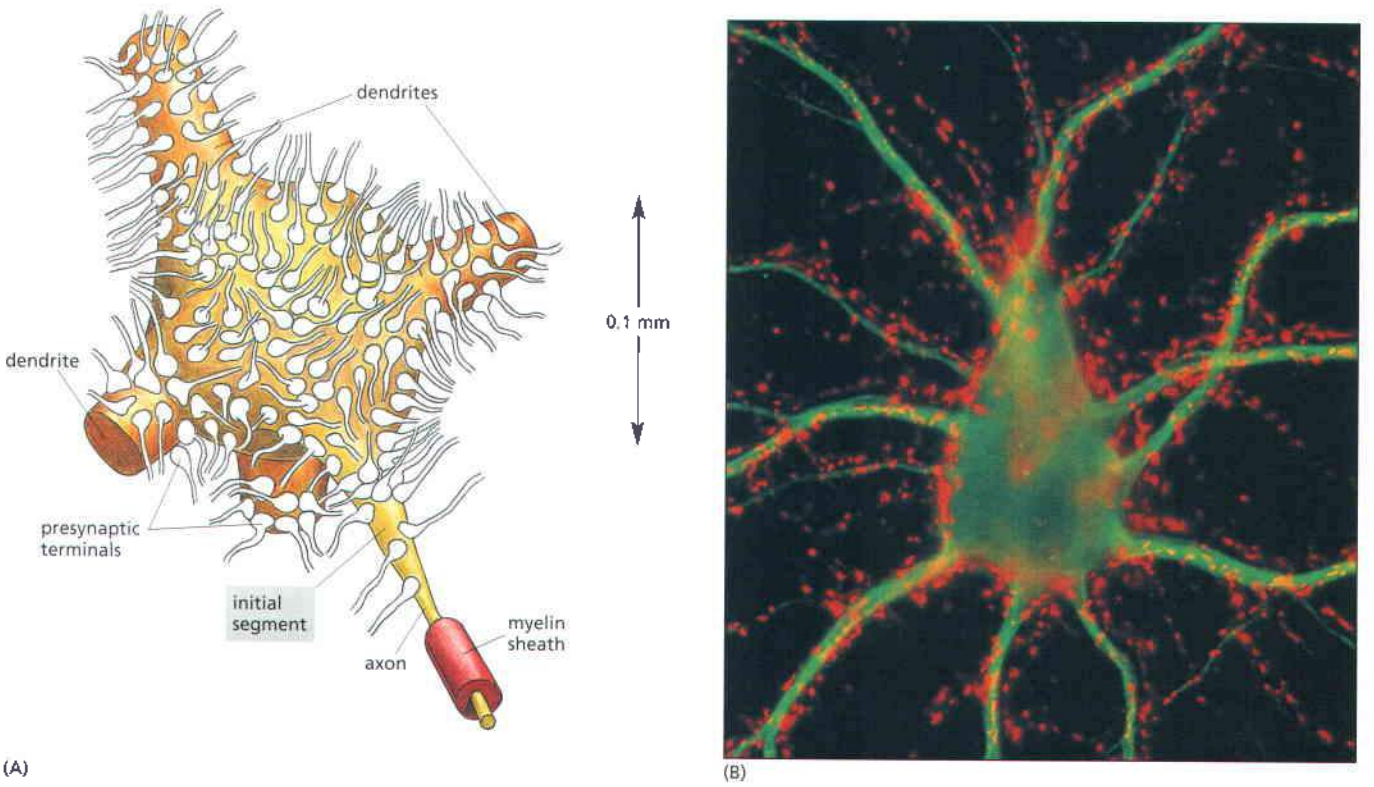


Figure 11–40 A motor neuron cell body in the spinal cord. (A) Many thousands of nerve terminals synapse on the cell body and dendrites. These deliver signals from other parts of the organism to control the firing of action potentials along the single axon of this large cell. (B) Micrograph showing a nerve cell body and its dendrites stained with a fluorescent antibody that recognizes a cytoskeletal protein *(green)*. Thousands of axon terminals *(red)* from other nerve cells (not visible) make synapses on the cell body and dendrites; they are stained with a fluorescent antibody that recognizes a protein in synaptic vesicles. (B, courtesy of Olaf Mundigl and Pietro de Camilli.)

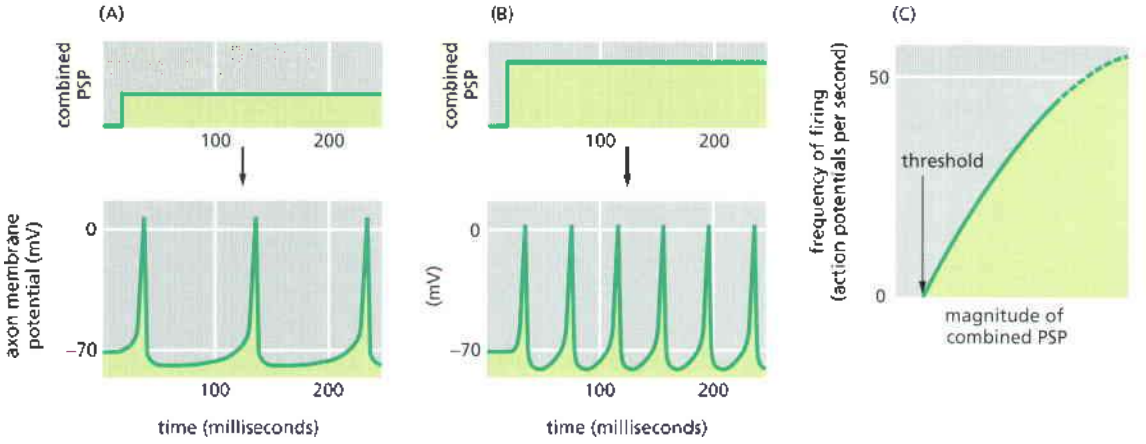


Figure 11–41 The magnitude of the combined postsynaptic potential (PSP) is reflected in the frequency of firing of action potentials. When successive action potentials arrive at the same synapse, each PSP produced adds to the preceding one to produce a larger combined PSP. A comparison of (A) and (B) shows how the firing frequency of an axon increases with an increase in the combined PSP, while (C) summarizes the general relationship.

trigger an action potential. Instead, each incoming signal is reflected in a local PSP of graded magnitude, which decreases with distance from the site of the synapse. If signals arrive simultaneously at several synapses in the same region of the dendritic tree, the total PSP in that neighborhood will be roughly the sum of the individual PSPs, with inhibitory PSPs making a negative contribution to the total. The PSPs from each neighborhood spread passively and converge on the cell body. For long-distance transmission, the combined magnitude of the PSP is then translated, or *encoded*, into the *frequency* of firing of action potentials (**Figure 11–41**). This encoding is achieved by a special set of gated ion channels that are present at high density at the base of the axon, adjacent to the cell body, in a region known as the *initial segment*, or *axon hillock* (see Figure 11–40).

Neuronal Computation Requires a Combination of at Least Three Kinds of K^+ Channels

We have seen that the intensity of stimulation a neuron receives is encoded for long-distance transmission by the frequency of action potentials that the neuron fires: the stronger the stimulation, the higher the frequency of action potentials. Action potentials are initiated at the **initial segment**, a unique region of each neuron with plentiful voltage-gated Na^+ channels. But to perform its special function of encoding, the membrane of the initial segment also contains at least four other classes of ion channels—three selective for K^+ and one selective for Ca^{2+}. The three varieties of K^+ channels have different properties; we shall refer to them as *delayed*, *rapidly inactivating*, and *Ca^{2+}-activated K^+ channels*.

To understand the need for multiple types of channels, consider first what would happen if the only voltage-gated ion channels present in the nerve cell were the Na^+ channels. Below a certain threshold level of synaptic stimulation, the depolarization of the initial segment membrane would be insufficient to trigger an action potential. With gradually increasing stimulation, the threshold would be crossed, the Na^+ channels would open, and an action potential would fire. The action potential would be terminated in the usual way by inactivation of the Na^+ channels. Before another action potential could fire, these channels would have to recover from their inactivation. But that would require a return of the membrane voltage to a very negative value, which would not occur as long as the strong depolarizing stimulus (from PSPs) was maintained. An additional channel type is needed, therefore, to repolarize the membrane after each action potential to prepare the cell to fire again.

The **delayed K^+ channels** perform this task, as discussed previously in relation to the propagation of the action potential (see p. 677). They are voltage-gated, but because of their slower kinetics they open only during the falling phase

of the action potential, when the Na^+ channels are inactive. Their opening permits an efflux of K^+ that drives the membrane back toward the K^+ equilibrium potential, which is so negative that the Na^+ channels rapidly recover from their inactivated state. Repolarization of the membrane also closes the delayed K^+ channels. The initial segment is now reset so that the depolarizing stimulus from synaptic inputs can fire another action potential. In this way, sustained stimulation of the dendrites and cell body leads to repetitive firing of the axon.

Repetitive firing in itself, however, is not enough. The frequency of the firing has to reflect the intensity of the stimulation, and a simple system of Na^+ channels and delayed K^+ channels is inadequate for this purpose. Below a certain threshold level of steady stimulation, the cell will not fire at all; above that threshold level, it will abruptly begin to fire at a relatively rapid rate. The **rapidly inactivating K^+ channels** solve the problem. These, too, are voltage-gated and open when the membrane is depolarized, but their specific voltage sensitivity and kinetics of inactivation are such that they act to reduce the rate of firing at levels of stimulation that are only just above the threshold required for firing. Thus, they remove the discontinuity in the relationship between the firing rate and the intensity of stimulation. The result is a firing rate that is proportional to the strength of the depolarizing stimulus over a very broad range (see Figure 11–41C).

The process of encoding is usually further modulated by the two other types of ion channels in the initial segment that were mentioned at the outset, namely voltage-gated Ca^{2+} channels and Ca^{2+}-activated K^+ channels. They act together to decrease the response of the cell to an unchanging, prolonged stimulation—a process called **adaptation**. These Ca^{2+} channels are similar to the Ca^{2+} channels that mediate the release of neurotransmitter from presynaptic axon terminals; they open when an action potential fires, transiently allowing Ca^{2+} into the initial segment.

The **Ca^{2+}-activated K^+ channel** is both structurally and functionally different from any of the channel types described earlier. It opens in response to a raised concentration of Ca^{2+} at the cytoplasmic face of the nerve cell membrane. Suppose we apply a strong depolarizing stimulus for a long time, triggering a long train of action potentials. Each action potential permits a brief influx of Ca^{2+} through the voltage-gated Ca^{2+} channels, so that the intracellular Ca^{2+} concentration gradually builds up to a level high enough to open the Ca^{2+}-activated K^+ channels. Because the resulting increased permeability of the membrane to K^+ makes the membrane harder to depolarize, it increases the delay between one action potential and the next. In this way, a neuron that is stimulated continuously for a prolonged period becomes gradually less responsive to the constant stimulus.

Such adaptation, which can also occur by other mechanisms, allows a neuron—indeed, the nervous system generally—to react sensitively to *change*, even against a high background level of steady stimulation. It is one of the strategies that help us, for example, to feel a light touch on the shoulder and yet ignore the constant pressure of our clothing. We discuss adaptation as a general feature in cell signaling processes in more detail in Chapter 15.

Other neurons do different computations, reacting to their synaptic inputs in myriad ways, reflecting the different assortments of members of the various ion channel families that reside in their membranes. There are several hundred genes that code for ion channels in the human genome, with over 150 encoding voltage-gated channels alone. Further complexity is introduced by alternative splicing of RNAs and assembling channels from different combinations of diverse subunits. The multiplicity of ion channels evidently allows for many different types of neurons, the electrical behavior of which is specifically tuned to the particular tasks that they must perform.

One of the crucial properties of the nervous system is its ability to learn and remember, which seems to depend largely on long-term changes in specific synapses. We end this chapter by considering a remarkable type of ion channel that is thought to have a special role in some forms of learning and memory. It is located at many synapses in the central nervous system, where it is gated by both voltage and the excitatory neurotransmitter glutamate. It is also the site of action of the psychoactive drug phencyclidine, or angel dust.

Long-Term Potentiation (LTP) in the Mammalian Hippocampus Depends on Ca^{2+} Entry Through NMDA-Receptor Channels

Practically all animals can learn, but mammals seem to learn exceptionally well (or so we like to think). In a mammal's brain, the region called the *hippocampus* has a special role in learning. When it is destroyed on both sides of the brain, the ability to form new memories is largely lost, although previous long-established memories remain. Correspondingly, some synapses in the hippocampus show marked functional alterations with repeated use: whereas occasional single action potentials in the presynaptic cells leave no lasting trace, a short burst of repetitive firing causes **long-term potentiation (LTP)**, such that subsequent single action potentials in the presynaptic cells evoke a greatly enhanced response in the postsynaptic cells. The effect lasts hours, days, or weeks, according to the number and intensity of the bursts of repetitive firing. Only the synapses that were activated exhibit LTP; synapses that have remained quiet on the same postsynaptic cell are not affected. However, while the cell is receiving a burst of repetitive stimulation via one set of synapses, if a single action potential is delivered at *another* synapse on its surface, that latter synapse also will undergo LTP, even though a single action potential delivered there at another time would leave no such lasting trace.

The underlying rule in such synapses seems to be that *LTP occurs on any occasion when a presynaptic cell fires (once or more) at a time when the post-synaptic membrane is strongly depolarized* (either through recent repetitive firing of the same presynaptic cell or by other means). This rule reflects the behavior of a particular class of ion channels in the postsynaptic membrane. Glutamate is the main excitatory neurotransmitter in the mammalian central nervous system, and glutamate-gated ion channels are the most common of all transmitter-gated channels in the brain. In the hippocampus, as elsewhere, most of the depolarizing current responsible for excitatory PSPs is carried by glutamate-gated ion channels, called **AMPA receptors**, that operate in the standard way. But the current has, in addition, a second and more intriguing component, which is mediated by a separate subclass of glutamate-gated ion channels known as **NMDA receptors**, so named because they are selectively activated by the artificial glutamate analog N-methyl-D-aspartate. The NMDA-receptor channels are doubly gated, opening only when two conditions are satisfied simultaneously: glutamate must be bound to the receptor, and the membrane must be strongly depolarized. The second condition is required for releasing the Mg^{2+} that normally blocks the resting channel. This means that NMDA receptors are normally activated only when AMPA receptors are activated as well and depolarize the membrane. The NMDA receptors are critical for LTP. When they are selectively blocked with a specific inhibitor, or in transgenic animals in which the gene has been knocked out, LTP does not occur, even though ordinary synaptic transmission continues. Such animals exhibit specific deficits in their learning abilities but behave almost normally otherwise.

How do NMDA receptors mediate such a remarkable effect? The answer is that these channels, when open, are highly permeable to Ca^{2+}, which acts as an intracellular mediator in the postsynaptic cell, triggering a cascade of changes that are responsible for LTP. Thus, LTP is prevented when Ca^{2+} levels are held artificially low in the postsynaptic cell by injecting the Ca^{2+} chelator EGTA into it, and LTP can be induced by artificially raising intracellular Ca^{2+} levels. Among the long-term changes that increase the sensitivity of the postsynaptic cell to glutamate is the insertion of new AMPA receptors into the plasma membrane (**Figure 11–42**). Evidence also indicates that changes can occur in the presynaptic cell as well, so that it releases more glutamate than normal when it is activated subsequently.

If synapses expressed only LTP they would quickly become saturated and, thus, be of limited value as an information storage device. In fact, synapses also exhibit **long-term depression (LTD)**, which surprisingly also requires NMDA receptor activation and a rise in Ca^{2+}. How does Ca^{2+} trigger opposite effects at the same synapse? It turns out that this bidirectional control of synaptic strength depends on the magnitude of the rise in Ca^{2+}: high Ca^{2+} levels activate protein kinases and LTP, whereas modest Ca^{2+} levels activate protein phosphatases and LTD.

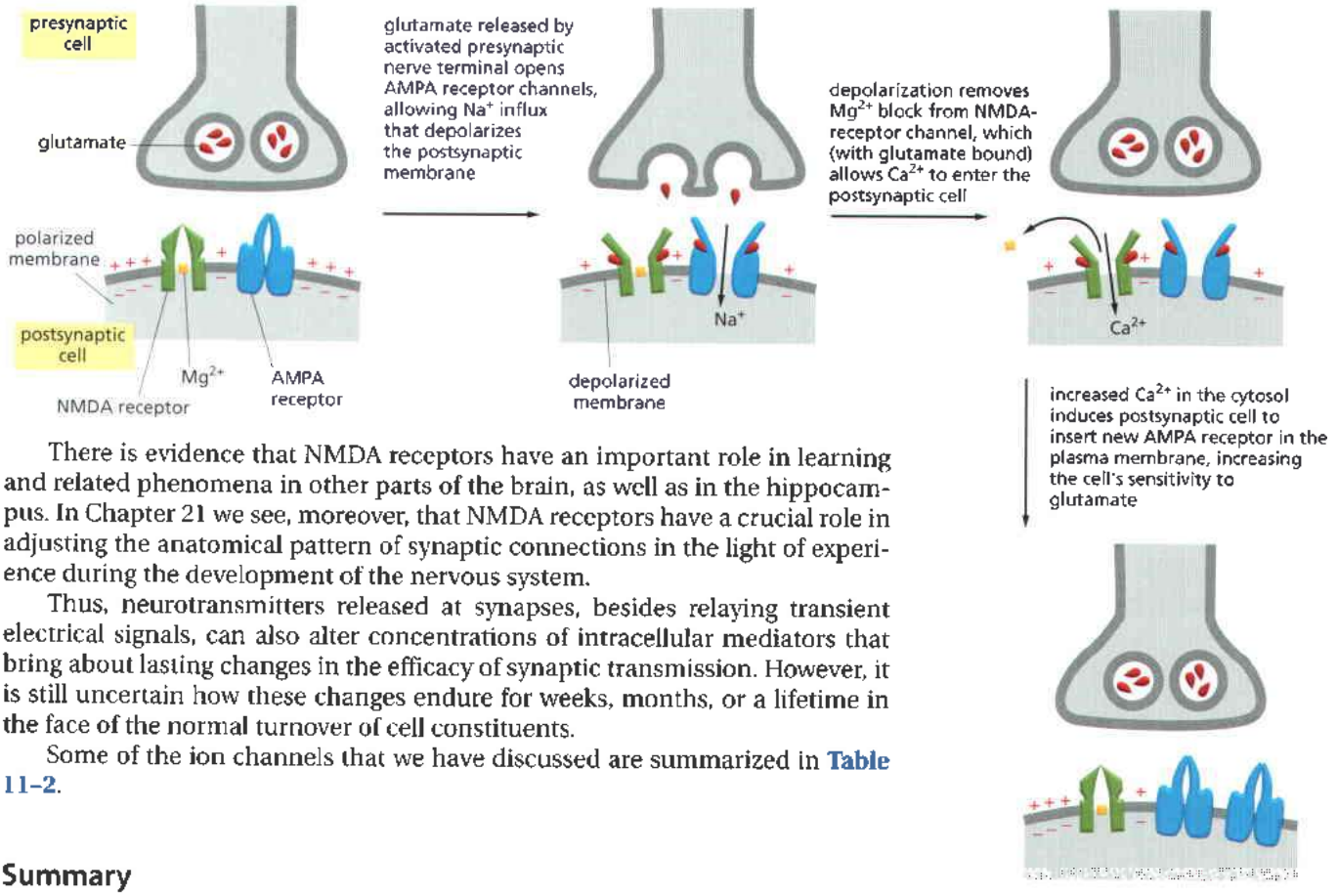


Figure 11–42 **The signaling events in long-term potentiation.** Although not shown, evidence suggests that changes can also occur in the presynaptic nerve terminals in LTP, which may be stimulated by retrograde signals from the postsynaptic cell.

There is evidence that NMDA receptors have an important role in learning and related phenomena in other parts of the brain, as well as in the hippocampus. In Chapter 21 we see, moreover, that NMDA receptors have a crucial role in adjusting the anatomical pattern of synaptic connections in the light of experience during the development of the nervous system.

Thus, neurotransmitters released at synapses, besides relaying transient electrical signals, can also alter concentrations of intracellular mediators that bring about lasting changes in the efficacy of synaptic transmission. However, it is still uncertain how these changes endure for weeks, months, or a lifetime in the face of the normal turnover of cell constituents.

Some of the ion channels that we have discussed are summarized in **Table 11–2**.

Summary

Ion channels form aqueous pores across the lipid bilayer and allow inorganic ions of appropriate size and charge to cross the membrane down their electrochemical gradients at rates about 1000 times greater than those achieved by any known transporter. The channels are "gated" and usually open transiently in response to a specific perturbation in the membrane, such as a change in membrane potential (voltage-gated channels) or the binding of a neurotransmitter (transmitter-gated channels).

K^+-selective leak channels have an important role in determining the resting membrane potential across the plasma membrane in most animal cells. Voltage-gated cation channels are responsible for the generation of self-amplifying action potentials in electrically excitable cells, such as neurons and skeletal muscle cells. Transmitter-gated ion channels convert chemical signals to electrical signals at chemical synapses. Excitatory neurotransmitters, such as acetylcholine and glutamate, open transmitter-gated cation channels and thereby depolarize the postsynaptic membrane toward the threshold level for firing an action potential. Inhibitory neurotransmitters, such as GABA and glycine, open transmitter-gated Cl^- or K^+ channels and thereby suppress firing by keeping the postsynaptic membrane polarized. A subclass of glutamate-gated ion channels, called NMDA-receptor channels, is highly permeable to Ca^{2+}, which can trigger the long-term changes in synapses such as LTP and LTD that are thought to be involved in some forms of learning and memory.

Table 11–2 Some Ion Channel Families

CHANNEL TYPE	REPRESENTATIVE EXAMPLE	
Voltage-gated cation channels	voltage-gated Na^+ channels voltage-gated K^+ channels (including delayed and early) voltage-gated Ca^{2+} channels	
Transmitter-gated ion channels	acetylcholine-gated cation channels glutamate-gated Ca^{2+} channels serotonin-gated cation channels	excitatory
	GABA-gated Cl^- channels glycine-gated Cl^- channels	inhibitory

Ion channels work together in complex ways to control the behavior of electrically excitable cells. A typical neuron, for example, receives thousands of excitatory and inhibitory inputs, which combine by spatial and temporal summation to produce a postsynaptic potential (PSP) in the cell body. The magnitude of the PSP is translated into the rate of firing of action potentials by a mixture of cation channels in the membrane of the initial segment.

PROBLEMS

Which statements are true? Explain why or why not.

11–1 Transport by transporters can be either active or passive, whereas transport by channels is always passive.

11–2 Transporters saturate at high concentrations of the transported molecule when all their binding sites are occupied; channels, on the other hand, do not bind the ions they transport and thus the flux of ions through a channel does not saturate.

11–3 The membrane potential arises from movements of charge that leave ion concentrations practically unaffected, causing only a very slight discrepancy in the number of positive and negative ions on the two sides of the membrane.

Discuss the following problems.

11–4 Order Ca^{2+}, CO_2, ethanol, glucose, RNA, and H_2O according to their ability to diffuse through a lipid bilayer, beginning with the one that crosses the bilayer most readily. Explain your order.

11–5 How is it possible for some molecules to be at equilibrium across a biological membrane and yet not be at the same concentration on both sides?

11–6 Ion transporters are "linked" together—not physically, but as a consequence of their actions. For example, cells can raise their intracellular pH, when it becomes too acidic, by exchanging external Na^+ for internal H^+, using a Na^+–H^+ antiporter. The change in internal Na^+ is then redressed using the Na^+-K^+ pump.

A. Can these two transporters, operating together, normalize both the H^+ and the Na^+ concentrations inside the cell?

B. Does the linked action of these two pumps cause imbalances in either the K^+ concentration or the membrane potential? Why or why not?

11–7 Microvilli increase the surface area of intestinal cells, providing more efficient absorption of nutrients. Microvilli are shown in profile and cross section in **Figure Q11–1**. From the dimensions given in the figure, estimate the increase in surface area that microvilli provide (for the portion of the plasma membrane in contact with the lumen of the gut) relative to the corresponding surface of a cell with a "flat" plasma membrane.

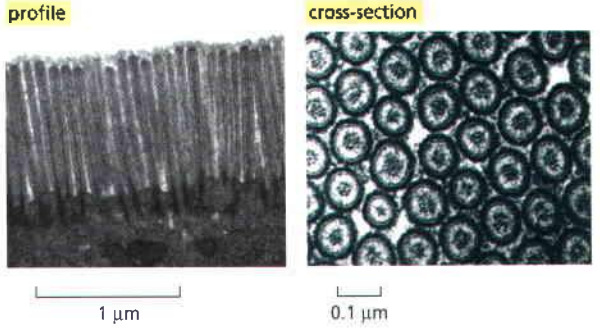


Figure Q11–1 Microvilli of intestinal epithelial cells in profile and cross section (Problem 11–7). (Left panel, from Rippel Electron Microscope Facility, Dartmouth College; Right panel, from David Burgess.)

11–8 According to Newton's laws of motion, an ion exposed to an electric field in a vacuum would experience a constant acceleration from the electric driving force, just as a falling body in a vacuum constantly accelerates due to gravity. In water, however, an ion moves at constant velocity in an electric field. Why do you suppose that is?

11–9 The "ball-and-chain" model for the rapid inactivation of voltage-gated K^+ channels has been elegantly confirmed for the *shaker* K^+ channel from *Drosophila melanogaster*. (The *shaker* K^+ channel in *Drosophila* is named after a mutant form that causes excitable behavior—even anesthetized flies keep twitching.) Deletion of the N-terminal amino acids from the normal *shaker* channel gives rise to a channel that opens in response to membrane depolarization, but stays open instead of rapidly closing as the normal channel does. A peptide (MAAVAGLYGLGEDRQHRKKQ) that corresponds to the deleted N-terminus can inactivate the open channel at 100 μM.

Is the concentration of free peptide (100 μM) that is required to inactivate the defective K^+ channel anywhere near the normal local concentration of the tethered ball on a normal channel? Assume that the tethered ball can explore a hemisphere [volume = $(2/3)\pi r^3$] with a radius of 21.4 nm, the length of the polypeptide "chain" (**Figure Q11–2**). Calculate the concentration for one ball in this hemisphere. How does that value compare with the concentration of free peptide needed to inactivate the channel?

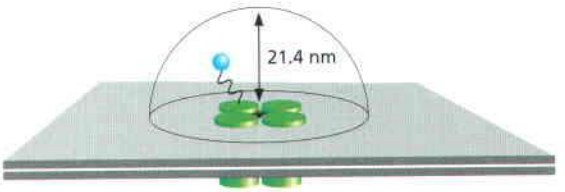


Figure Q11–2 A "ball" tethered by a "chain" to a voltage-gated K^+ channel (Problem 11–9).

11–10 The squid giant axon occupies a unique position in the history of our understanding of cell membrane potentials and nerve action. When an electrode is stuck into an intact giant axon, the membrane potential registers –70 mV. When the axon, suspended in a bath of seawater, is stimulated to conduct a nerve impulse, the membrane potential changes transiently from –70 mV to +40 mV.

For univalent ions and at 20°C (293 K), the Nernst equation reduces to

$$V = 58 \text{ mV} \times \log (C_o/C_i)$$

where C_o and C_i are the concentrations outside and inside, respectively.

Using this equation, calculate the potential across the resting membrane (1) assuming that it is due solely to K^+ and (2) assuming that it is due solely to Na^+. (The Na^+ and K^+ concentrations in axon cytoplasm and in seawater are given in Table Q11–1.) Which calculation is closer to the measured resting potential? Which calculation is closer to the measured action potential? Explain why these assumptions approximate the measured resting and action potentials.

Table Q11–1 Ionic composition of seawater and of cytoplasm from the squid giant axon (Problem 11–10).

ION	CYTOPLASM	SEAWATER
Na^+	65 mM	430 mM
K^+	344 mM	9 mM

REFERENCES

General

Martonosi AN (ed) (1985) The enzymes of Biological membranes, vol 3: Membrane transport. 2nd edn. New York: Penum Press.

Stein WD (1990) Channels, carriers and pumps: An ntroduction to membrane transport. San Diego: Academic Press

Principles of Membrane Transport

Al-Awqati Q (1999) One hundred years of membrane permeability: does Overton still rule? *Nature Cell Biol* 1: E201–E202.

Forrest LR & Sansom MS (2000) Membrane simulations: bigger and better? *Curr Opin Struct Biol* 10:174–181.

Gouaux E and MacKinnon R (2005) Principles of selective ion transport in channels and pumps. *Science* 310:1461–1465.

Mitchell P (1977) Vectorial chemiosmotic processes. *Annu Rev Biochem* 46:996–1005.

Tanford C (1983) Mechanism of free energy coupling in active transport. *Annu Rev Biochem* 52:379–409.

Carrier Proteins and Active Membrane Transport

Almers W & Stirling C (1984) Distribution of transport proteins over animal cell membranes. *J Membr Biol* 77:169–186.

Baldwin SA & Henderson PJ (1989) Homologies between sugar transporters from eukaryotes and prokaryotes. *Annu Rev Physiol* 51:459–471.

Borst P & Elferink RO (2002) Mammalian ABC transporters in health and disease. *Annu Rev Biochem* 71:537–592.

Carafoli E & Brini M (2000) Calcium pumps: structural basis for and mechanism of calcium transmembrane transport. *Curr Opin Chem Biol* 4:152–160.

Dean M, Rzhetsky A et al (2001) The human ATP-binding cassette (ABC) transporter superfamily. *Genome Res* 11:1156–1166.

Doige CA & Ames GF (1993) ATP-dependent transport systems in bacteria and humans: relevance to cystic fibrosis and multidrug resistance. *Annu Rev Microbiol* 47:291–319.

Gadsby DC, Vergani P & Csanady L (2006) The ABC protein turned chloride channel whose failure causes cystic fibrosis. *Nature* 440:477–83.

Higgins CF (2007) Multiple molecular mechanisms for multidrug resistance transporters. *Nature* 446:749–57.

Kaback HR, Sahin-Toth M et al (2001) The kamikaze approach to membrane transport. *Nature Rev Mol Cell Biol* 2:610–620.

Kühlbrandt W (2004) Biology, structure and mechanism of P-type ATPases. *Nature Rev Mol Cell Biol* 5:282–295.

Lodish HF (1986) Anion-exchange and glucose transport proteins: structure, function, and distribution. *Harvey Lect* 82:19–46.

Pedersen PL & Carafoli E (1987) Ion motive ATPases. 1. Ubiquity, properties, and significance to cell function. *Trends Biochem Sci* 12:146–150.

Romero MF & Boron WF (1999) Electrogenic Na^+/HCO_3^- cotransporters: cloning and physiology. *Annu Rev Physiol* 61:699–723.

Saier MH, Jr (2000) Vectorial metabolism and the evolution of transport systems. *J Bacteriol* 182:5029–5035.

Scarborough GA (2003) Rethinking the P-type ATPase problem. *Trends Biochem Sci* 28:581–584.

Stein WD (2002) Cell volume homeostasis: ionic and nonionic mechanisms. The sodium pump in the emergence of animal cells. *Int Rev Cytol* 215:231–258.

Ion Channels and the Electrical Properties of Membranes

Armstrong C (1998) The vision of the pore. *Science* 280:56–57.

Choe S (2002) Potassium channel structures. *Nature Rev Neurosci* 3:115–21.

Choe S, Kreusch A & Pfaffinger PJ (1999) Towards the three-dimensional structure of voltage-gated potassium channels. *Trends Biochem Sci* 24:345–349.

Franks NP & Lieb WR (1994) Molecular and cellular mechanisms of general anaesthesia. *Nature* 367:607–614.

Greengard P (2001) The neurobiology of slow synaptic transmission. *Science* 294:1024–30.

Hille B (2001) Ionic Channels of Excitable Membranes, 3rd ed. Sunderland, MA: Sinauer.

Hucho F, Tsetlin VI & Machold J (1996) The emerging three-dimensional structure of a receptor. The nicotinic acetylcholine receptor. *Eur J Biochem* 239:539–557.

Hodgkin AL & Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117:500–544.

Hodgkin AL & Huxley AF (1952) Currents carried by sodium and potassium ions through the membrane of the giant axon of *Loligo*. *J Physiol* 116:449–472.

Jessell TM & Kandel ER (1993) Synaptic transmission: a bidirectional and self-modifiable form of cell–cell communication. *Cell* 72(Suppl):1–30.

Kandel ER, Schwartz JH & Jessell TM (2000) Principles of Neural Science, 4th ed. New York: McGraw-Hill.

Karlin A (2002) Emerging structure of the nicotinic acetylcholine receptors. *Nature Rev Neurosci* 3:102–114.

Katz B (1966) Nerve, Muscle and Synapse. New York: McGraw-Hill.

King LS, Kozono D & Agre P (2004) From structure to disease: the evolving tale of aquaporin biology. *Nature Rev Mol Cell Biol* 5:687–698.

MacKinnon R (2003) Potassium channels. *FEBS Lett* 555:62–65.

Malenka RC & Nicoll RA (1999) Long-term potentiation—a decade of progress? *Science* 285:1870–1874.

Moss SJ & Smart TG (2001) Constructing inhibitory synapses. *Nature Rev Neurosci* 2:240–250.

Neher E and Sakmann B (1992) The patch clamp technique. *Sci Am* 266:44–51.

Nicholls JG, Fuchs PA, Martin AR & Wallace BG (2000) From Neuron to Brain, 4th ed. Sunderland, MA: Sinauer.

Numa S (1987) A molecular view of neurotransmitter receptors and ionic channels. *Harvey Lect* 83:121–165.

Scannevin RH & Huganir RL (2000) Postsynaptic organization and regulation of excitatory synapses. *Nature Rev Neurosci* 1:133–141.

Seeburg PH (1993) The molecular biology of mammalian glutamate receptor channels. *Trends Neurosci* 16:359–365.

Snyder SH (1996) Drugs and the Brain. New York: WH Freeman/Scientific American Books.

Stevens CF (2004) Presynaptic function. *Curr Opin Neurobiol* 14:341–345.

Tsien RW, Lipscombe D, Madison DV et al (1988) Multiple types of neuronal calcium channels and their selective modulation. *Trends Neurosci* 11:431–438.

Unwin N (2003) Structure and action of the nicotinic acetylcholine receptor explored by electron microscopy. *FEBS Lett* 555:91–95.

Intracellular Compartments and Protein Sorting

12

Unlike a bacterium, which generally consists of a single intracellular compartment surrounded by a plasma membrane, a eucaryotic cell is elaborately subdivided into functionally distinct, membrane-enclosed compartments. Each compartment, or **organelle**, contains its own characteristic set of enzymes and other specialized molecules, and complex distribution systems transport specific products from one compartment to another. To understand the eucaryotic cell, it is essential to know how the cell creates and maintains these compartments, what occurs in each of them, and how molecules move between them.

Proteins confer upon each compartment its characteristic structural and functional properties. They catalyze the reactions that occur in each organelle and selectively transport small molecules into and out of its interior, or **lumen**. Proteins also serve as organelle-specific surface markers that direct new deliveries of proteins and lipids to the appropriate organelle.

An animal cell contains about 10 billion (10^{10}) protein molecules of perhaps 10,000 kinds, and the synthesis of almost all of them begins in the cytosol. Each newly synthesized protein is then delivered specifically to the cell compartment that requires it. The intracellular transport of proteins is the central theme of both this chapter and the next. By tracing the protein traffic from one compartment to another, one can begin to make sense of the otherwise bewildering maze of intracellular membranes.

In This Chapter

THE COMPARTMENTALIZATION OF CELLS 695

THE TRANSPORT OF MOLECULES BETWEEN THE NUCLEUS AND THE CYTOSOL 704

THE TRANSPORT OF PROTEINS INTO MITOCHONDRIA AND CHLOROPLASTS 713

PEROXISOMES 721

THE ENDOPLASMIC RETICULUM 723

THE COMPARTMENTALIZATION OF CELLS

In this brief overview of the compartments of the cell and the relationships between them, we organize the organelles conceptually into a small number of discrete families, discuss how proteins are directed to specific organelles, and explain how proteins cross organelle membranes.

All Eucaryotic Cells Have the Same Basic Set of Membrane-Enclosed Organelles

Many vital biochemical processes take place in or on membrane surfaces. Membrane-bound enzymes, for example, catalyze lipid metabolism, and oxidative phosphorylation and photosynthesis both require a membrane to couple the transport of H^+ to the synthesis of ATP. In addition to providing increased membrane area to host biochemical reactions, intracellular membrane systems form enclosed compartments that are separate from the cytosol, thus creating functionally specialized aqueous spaces within the cell. Because the lipid bilayer of organelle membranes is impermeable to most hydrophilic molecules, the membrane of each organelle must contain membrane transport proteins to import and export specific metabolites. Each organelle membrane must also have a mechanism for importing, and incorporating into the organelle, the specific proteins that make the organelle unique.

Figure 12–1 illustrates the major intracellular compartments common to eucaryotic cells. The *nucleus* contains the genome (aside from mitochondrial

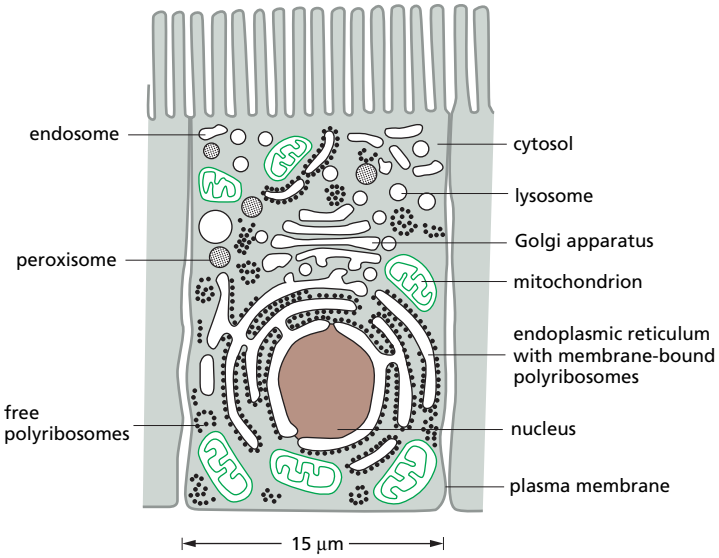


Figure 12–1 The major intracellular compartments of an animal cell. <ATCC> The cytosol *(gray),* endoplasmic reticulum, Golgi apparatus, nucleus, mitochondrion, endosome, lysosome, and peroxisome are distinct compartments isolated from the rest of the cell by at least one selectively permeable membrane.

and chloroplast DNA) and it is the principal site of DNA and RNA synthesis. The surrounding **cytoplasm** consists of the cytosol and the cytoplasmic organelles suspended in it. The **cytosol** constitutes a little more than half the total volume of the cell, and it is the site of protein synthesis and degradation. It also performs most of the cell's intermediary metabolism—that is, the many reactions that degrade some small molecules and synthesize others to provide the building blocks for macromolecules (discussed in Chapter 2).

About half the total area of membrane in a eucaryotic cell encloses the labyrinthine spaces of the *endoplasmic reticulum (ER)*. The *rough ER* has many ribosomes bound to its cytosolic surface; these synthesize both soluble and integral membrane proteins, most of which are destined either for secretion to the cell exterior or for other organelles. We shall see that, whereas proteins are transported into other organelles only after their synthesis is complete, they are transported into the ER as they are synthesized. This explains why the ER membrane is unique in having ribosomes tethered to it. The ER also produces most of the lipid for the rest of the cell and functions as a store for Ca^{2+} ions. Regions of the ER that lack bound ribosomes are called *smooth ER*. The ER sends many of its proteins and lipids to the *Golgi apparatus*, which consists of organized stacks of disclike compartments called *Golgi cisternae*. The Golgi apparatus receives lipids and proteins from the ER and dispatches them to various destinations, usually covalently modifying them *en route*.

Mitochondria and (in plants) *chloroplasts* generate most of the ATP that cells use to drive reactions requiring an input of free energy; chloroplasts are a specialized version of *plastids,* which can also have other functions in plant cells, such as the storage of food or pigment molecules. *Lysosomes* contain digestive enzymes that degrade defunct intracellular organelles, as well as macromolecules and particles taken in from outside the cell by endocytosis. On their way to lysosomes, endocytosed material must first pass through a series of organelles called *endosomes*. Finally, *peroxisomes* are small vesicular compartments that contain enzymes used in various oxidation reactions.

In general, each membrane-enclosed organelle performs the same set of basic functions in all cell types. But to serve the specialized functions of cells, these organelles vary in abundance and can have additional properties that differ from cell type to cell type.

On average, the membrane-enclosed compartments together occupy nearly half the volume of a cell (**Table 12–1**), and a large amount of intracellular membrane is required to make them all. In liver and pancreatic cells, for example, the endoplasmic reticulum has a total membrane surface area that is, respectively, 25 times and 12 times that of the plasma membrane (**Table 12–2**). In terms of its area and mass, the plasma membrane is only a minor membrane in most eucaryotic cells, and organelles are packed tightly in the cytosol (**Figure 12–2**).

Table 12–1 Relative Volumes Occupied by the Major Intracellular Compartments in a Liver Cell (Hepatocyte)

INTRACELLULAR COMPARTMENT	PERCENTAGE OF TOTAL CELL VOLUME
Cytosol	54
Mitochondria	22
Rough ER cisternae	9
Smooth ER cisternae plus Golgi cisternae	6
Nucleus	6
Peroxisomes	1
Lysosomes	1
Endosomes	1

Membrane-enclosed organelles often have characteristic positions in the cytosol. In most cells, for example, the Golgi apparatus is located close to the nucleus, whereas the network of ER tubules extends from the nucleus throughout the entire cytosol. These characteristic distributions depend on interactions of the organelles with the cytoskeleton. The localization of both the ER and the Golgi apparatus, for instance, depends on an intact microtubule array; if the microtubules are experimentally depolymerized with a drug, the Golgi apparatus fragments and disperses throughout the cell, and the ER network collapses toward the cell center (discussed in Chapter 16).

Evolutionary Origins Explain the Topological Relationships of Organelles

To understand the relationships between the compartments of the cell, it is helpful to consider how they might have evolved. The precursors of the first eucaryotic cells are thought to have been simple organisms that resembled bacteria, which generally have a plasma membrane but no internal membranes. The plasma membrane in such cells therefore provides all membrane-dependent functions, including the pumping of ions, ATP synthesis, protein secretion, and lipid synthesis. Typical present-day eucaryotic cells are 10–30 times larger in linear dimension and 1000–10,000 times greater in volume than a typical bacterium

Table 12–2 Relative Amounts of Membrane Types in Two Kinds of Eucaryotic Cells

MEMBRANE TYPE	PERCENTAGE OF TOTAL CELL MEMBRANE	
	LIVER HEPATOCYTE*	PANCREATIC EXOCRINE CELL*
Plasma membrane	2	5
Rough ER membrane	35	60
Smooth ER membrane	16	<1
Golgi apparatus membrane	7	10
Mitochondria		
Outer membrane	7	4
Inner membrane	32	17
Nucleus		
Inner membrane	0.2	0.7
Secretory vesicle membrane	not determined	3
Lysosome membrane	0.4	not determined
Peroxisome membrane	0.4	not determined
Endosome membrane	0.4	not determined

*These two cells are of very different sizes: the average hepatocyte has a volume of about 5000 μm^3 compared with 1000 μm^3 for the pancreatic exocrine cell. Total cell membrane areas are estimated at about 110,000 μm^2 and 13,000 μm^2, respectively.

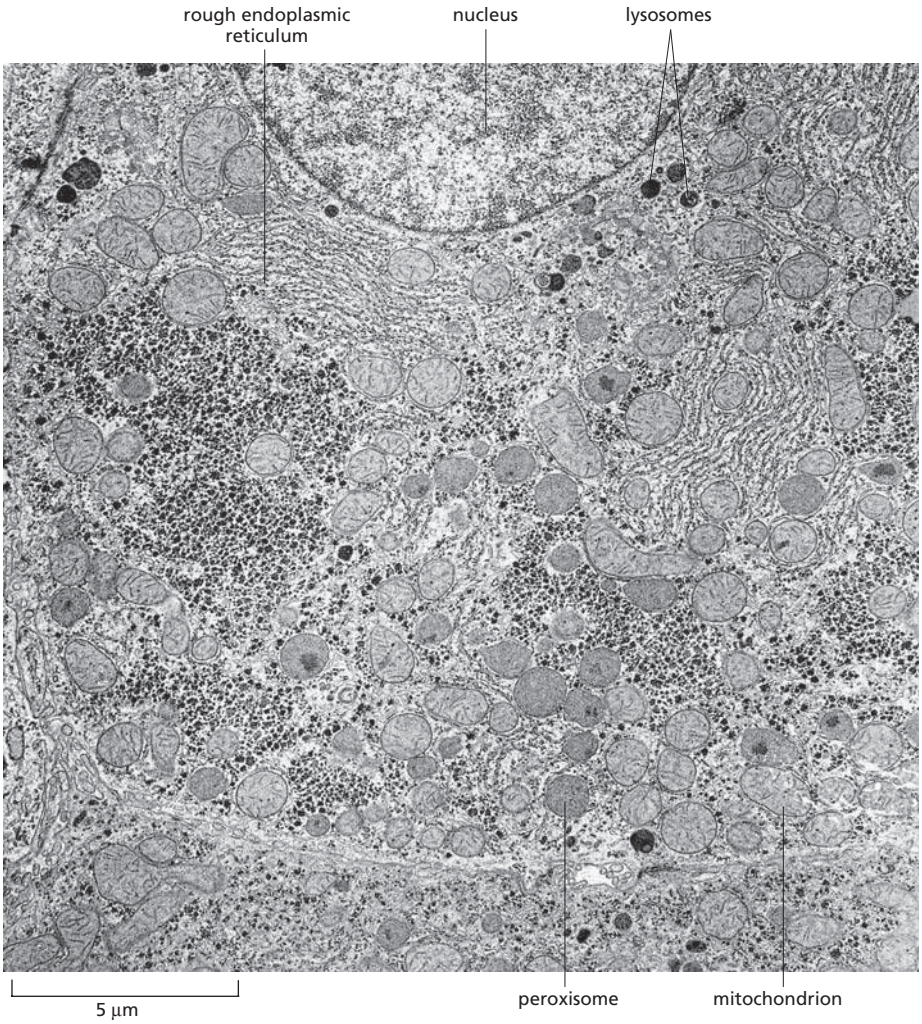


Figure 12–2 An electron micrograph of part of a liver cell seen in cross section. Examples of most of the major intracellular compartments are indicated. (Courtesy of Daniel S. Friend.)

such as *E. coli*. The profusion of internal membranes can be seen, in part, as an adaptation to this increase in size: the eucaryotic cell has a much smaller ratio of surface area to volume, and its plasma membrane therefore presumably has too small an area to sustain the many vital functions that membranes perform. The extensive internal membrane systems of a eucaryotic cell alleviate this problem.

The evolution of internal membranes evidently accompanied the specialization of membrane function. Consider, for example, the generation of *thylakoid vesicles* in chloroplasts. These vesicles form when chloroplasts in green leaf cells develop from proplastids, small precursor organelles that are present in all immature plant cells. These organelles are surrounded by a double membrane and develop according to the needs of the differentiated cells: they turn into chloroplasts in leaf cells, for example, but into organelles that store starch, fat, or pigments in other cell types (**Figure 12–3**A). When they convert into chloroplasts, specialized membrane patches form, invaginate, and pinch off from the inner membrane of the proplastid. The vesicles that result form a new specialized compartment, the *thylakoid,* which harbors all of the chloroplast's photosynthetic machinery (Figure 12–3B).

Other compartments in eucaryotic cells may have originated in a conceptually similar way (**Figure 12–4**). The invagination and pinching off of specialized intracellular membrane structures from the plasma membrane creates organelles with an interior that is topologically equivalent to the exterior of the cell. We shall see that this topological relationship holds for all of the organelles involved in the secretory and endocytic pathways, including the ER, Golgi apparatus, endosomes, and lysosomes. We can therefore think of all of these organelles as members of the same family. As we discuss in detail in the next chapter, their interiors communicate extensively with one another and with the

outside of the cell via *transport vesicles*, which bud off from one organelle and fuse with another (**Figure 12–5**).

As described in Chapter 14, mitochondria and plastids differ from the other membrane-enclosed organelles because they contain their own genomes. The nature of these genomes, and the close resemblance of the proteins in these organelles to those in some present-day bacteria, strongly suggest that mitochondria and plastids evolved from bacteria that were engulfed by other cells with which they initially lived in symbiosis (discussed in Chapters 1 and 14). According to the hypothetical scheme shown in Figure 12–4B, the inner membrane of mitochondria and plastids corresponds to the original plasma membrane of the bacterium, while the lumen of these organelles evolved from the bacterial cytosol. As we might expect from such an endocytic origin, these two organelles are surrounded by a double membrane, and they remain isolated from the extensive vesicular traffic that connects the interiors of most of the other membrane-enclosed organelles to each other and to the outside of the cell.

The evolutionary scheme described above groups the intracellular compartments in eucaryotic cells into four distinct families: (1) the nucleus and the cytosol, which communicate with each other through *nuclear pore complexes* and are thus topologically continuous (although functionally distinct); (2) all organelles that function in the secretory and endocytic pathways—including the ER, Golgi apparatus, endosomes, and lysosomes, the numerous classes of transport intermediates such as transport vesicles that move between them, and possibly peroxisomes; (3) the mitochondria; and (4) the plastids (in plants only).

Proteins Can Move Between Compartments in Different Ways

The synthesis of all proteins begins on ribosomes in the cytosol, except for the few that are synthesized on the ribosomes of mitochondria and plastids. Their subsequent fate depends on their amino acid sequence, which can contain **sorting signals** that direct their delivery to locations outside the cytosol. Most proteins do not have a sorting signal and consequently remain in the cytosol as permanent residents. Many others, however, have specific sorting signals that direct their transport from the cytosol into the nucleus, the ER, mitochondria, plastids, or peroxisomes; sorting signals can also direct the transport of proteins from the ER to other destinations in the cell.

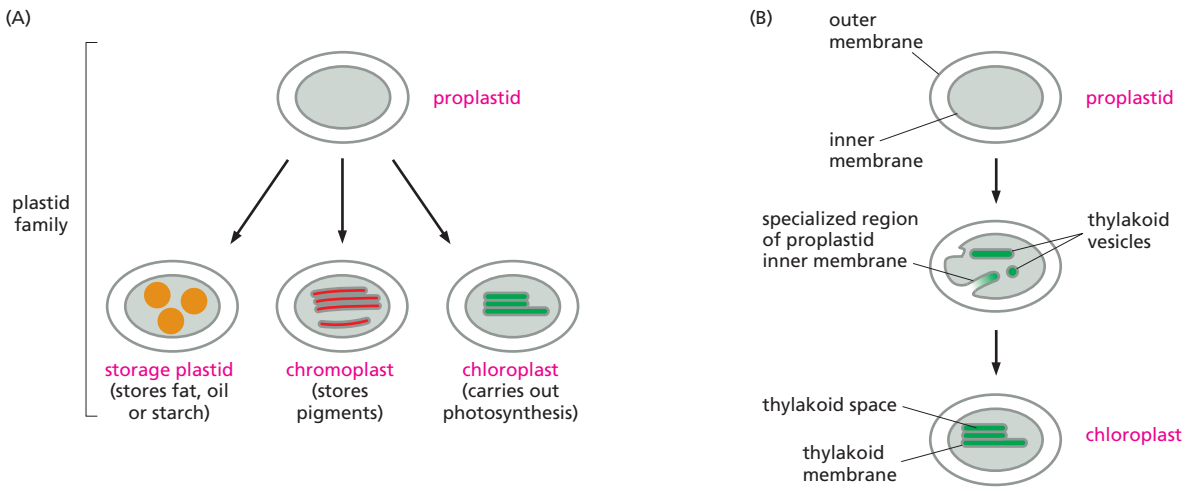


Figure 12–3 Development of plastids. (A) Proplastids are inherited with the cytoplasm of plant egg cells. As immature plant cells differentiate, the proplastids develop according to the needs of the specialized cell: they can become chloroplasts (in green leaf cells), storage plastids that accumulate starch (e.g., in potato tubers) or oil and lipid droplets (e.g., in fatty seeds), or chromoplasts that harbor pigments (e.g., in flower petals). (B) Development of the thylakoid. As chloroplasts develop, patches of specialized membrane in the proplastid inner membrane invaginate and pinch off to form thylakoid vesicles, which then develop into the mature thylakoid. The thylakoid membrane forms a separate compartment, the thylakoid space, which is structurally and functionally distinct from the rest of the chloroplast. Thylakoids can grow and divide autonomously as chloroplasts proliferate.

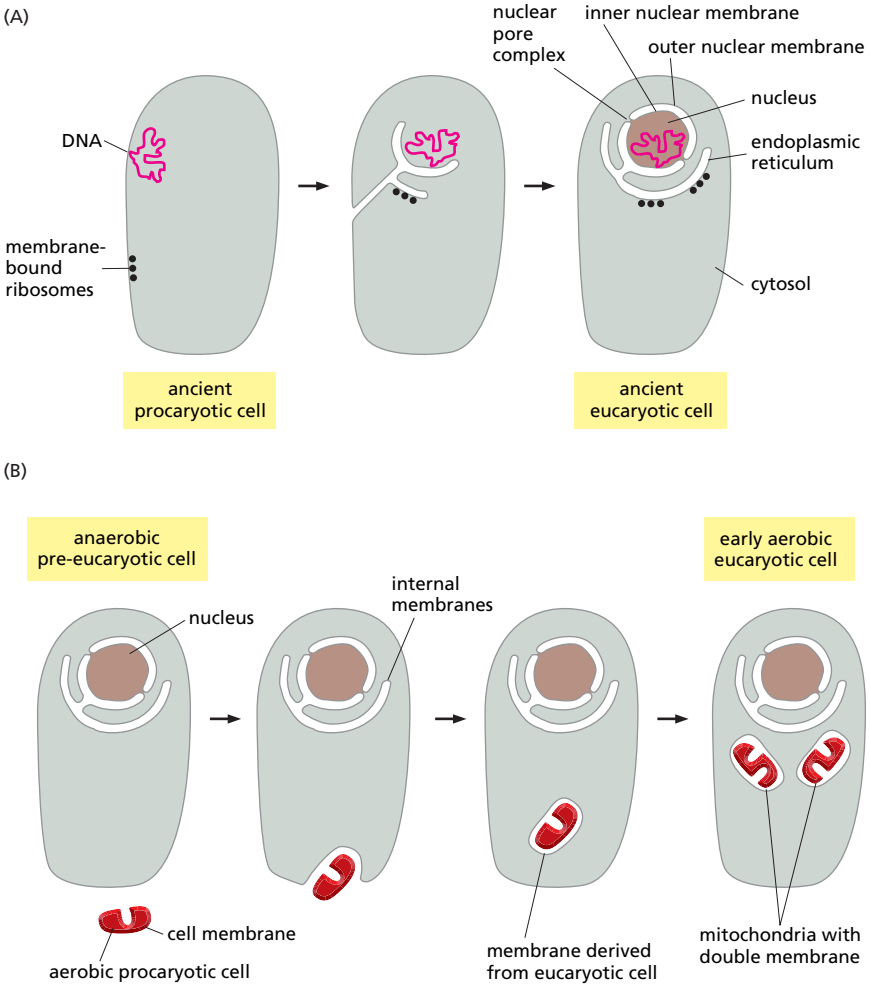


Figure 12–4 Hypothetical schemes for the evolutionary origins of some organelles. The origins of mitochondria, chloroplasts, ER, and the cell nucleus might explain the topological relationships of these compartments in eucaryotic cells.

(A) A possible pathway for the evolution of the cell nucleus and the ER. In some bacteria the single DNA molecule that comprises the bacterial chromosome is attached to an invagination of the plasma membrane. Such an invagination in a very ancient procaryotic cell could have rearranged to form an envelope around the DNA, while still allowing the DNA access to the cell cytosol (as is required for DNA to direct protein synthesis). This envelope is presumed to have eventually pinched off completely from the plasma membrane, producing a nuclear compartment surrounded by a double membrane.

As illustrated, gated passageways called nuclear pore complexes (NPCs) penetrate the nuclear envelope. Because it is surrounded by two membranes that are in continuity where they are penetrated by NPCs, the nuclear compartment is topologically equivalent to the cytosol; in fact, during mitosis the nuclear contents mix with the cytosol. The lumen of the ER is continuous with the space between the inner and outer nuclear membranes, and it is topologically equivalent to the extracellular space (see Figure 12–5).

(B) Mitochondria (and plastids) are thought to have originated when a bacterium was engulfed by a larger pre-eucaryotic cell. This could explain why they contain their own genomes and why the lumens of these organelles remain isolated from the membrane traffic that interconnects the lumens of many other intracellular compartments.

To understand the general principles by which sorting signals operate, it is important to distinguish three fundamentally different ways by which proteins move from one compartment to another. These three mechanisms are described below, and their sites of action in the cell are outlined in **Figure 12–6**. We discuss the first two mechanisms in this chapter, and the third *(green arrows* in Figure 12–6) in Chapter 13.

1. In **gated transport**, proteins move between the cytosol and the nucleus (which are topologically equivalent) through nuclear pore complexes in the nuclear envelope. The nuclear pore complexes function as selective gates that actively transport specific macromolecules and macromolecular assemblies, although they also allow free diffusion of smaller molecules.
2. In **transmembrane transport**, transmembrane *protein translocators* directly transport specific proteins across a membrane from the cytosol into a space that is topologically distinct. The transported protein molecule

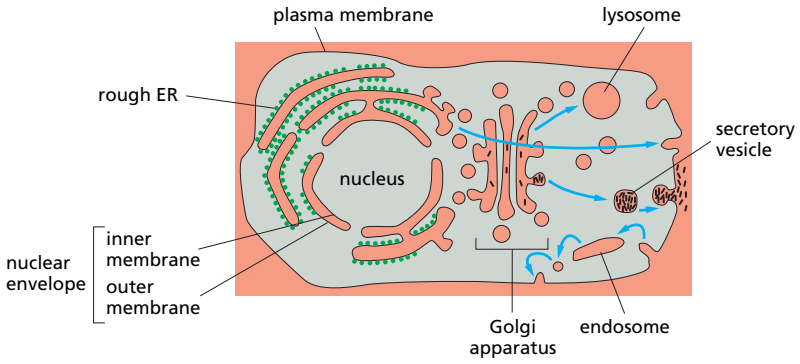


Figure 12–5 Topological relationships between compartments of the secretory and endocytic pathways in a eucaryotic cell. Topologically equivalent spaces are shown in *red*. In principle, cycles of membrane budding and fusion permit the lumen of any of these organelles to communicate with any other and with the cell exterior by means of transport vesicles. *Blue arrows* indicate the extensive outbound and inbound vesicular traffic (discussed in Chapter 13). Some organelles, most notably mitochondria and (in plant cells) plastids, do not take part in this communication and are isolated from the traffic between organelles shown here.

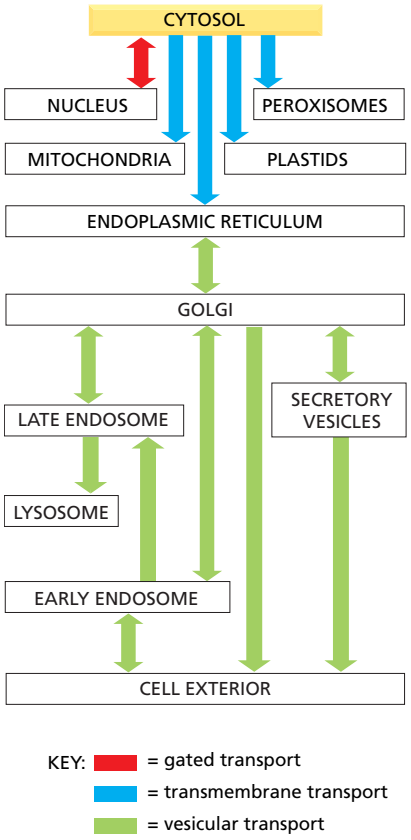


Figure 12–6 A simplified "roadmap" of protein traffic. Proteins can move from one compartment to another by gated transport *(red)*, transmembrane transport *(blue)*, or vesicular transport *(green)*. The sorting signals that direct a given protein's movement through the system, and thereby determine its eventual location in the cell, are contained in each protein's amino acid sequence. The journey begins with the synthesis of a protein on a ribosome in the cytosol and terminates when the protein reaches its final destination. At each intermediate station *(boxes)*, a decision is made as to whether the protein is to be retained in that compartment or transported further. In principle, a sorting signal could be required for either retention in or exit from a compartment.

We shall refer to this figure often as a guide in this chapter and the next, highlighting in color the particular pathway being discussed.

usually must unfold to snake through the translocator. The initial transport of selected proteins from the cytosol into the ER lumen or mitochondria, for example, occurs in this way.

3. In **vesicular transport**, membrane-enclosed transport intermediates—which may be small, spherical transport vesicles or larger, irregularly shaped organelle fragments—ferry proteins from one compartment to another. The transport vesicles and fragments become loaded with a cargo of molecules derived from the lumen of one compartment as they bud and pinch off from its membrane; they discharge their cargo into a second compartment by fusing with the membrane enclosing that compartment (**Figure 12–7**). The transfer of soluble proteins from the ER to the Golgi apparatus, for example, occurs in this way. Because the transported proteins do not cross a membrane, vesicular transport can move proteins only between compartments that are topologically equivalent (see Figure 12–5).

Each mode of protein transfer is usually guided by sorting signals in the transported protein, which are recognized by complementary *sorting receptors*. If a large protein is to be imported into the nucleus, for example, it must possess a sorting signal that receptor proteins recognize to guide it through the nuclear pore complex. If a protein is to be transferred directly across a membrane, it must possess a sorting signal that the membrane translocator recognizes. Likewise, if a protein is to be loaded into a certain type of vesicle or retained in certain organelles, a complementary receptor in the appropriate membrane must recognize its sorting signal.

Signal Sequences Direct Proteins to the Correct Cell Address

Most protein sorting signals reside in a stretch of amino acid sequence, typically 15–60 residues long. These **signal sequences** are often found at the N-terminus; specialized **signal peptidases** remove the signal sequence from the finished protein once the sorting process is complete. Signal sequences can also be internal stretches of amino acids, which remain part of the protein. In some cases, sorting signals are composed of multiple internal amino acid sequences that form a specific three-dimensional arrangement of atoms on the protein's surface, called a **signal patch**.

Each signal sequence specifies a particular destination in the cell. Proteins destined for initial transfer to the ER usually have a signal sequence at their N-terminus, which characteristically includes a sequence composed of about 5–10

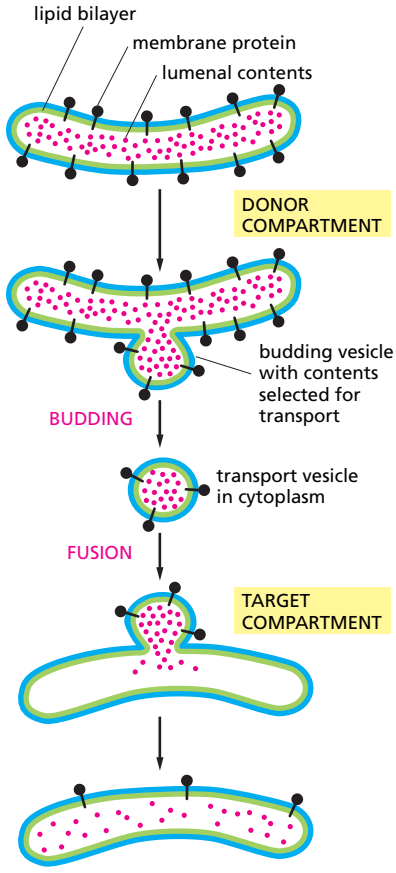


Figure 12–7 Vesicle budding and fusion during vesicular transport. Transport vesicles bud from one compartment (donor) and fuse with another (target) compartment. In the process, soluble components *(red dots)* are transferred from lumen to lumen. Note that membrane is also transferred and that the original orientation of both proteins and lipids in the donor-compartment membrane is preserved in the target-compartment membrane. Thus, membrane proteins retain their asymmetric orientation, with the same domains always facing the cytosol.

hydrophobic amino acids. Many of these proteins will in turn pass from the ER to the Golgi apparatus, but those with a specific signal sequence of four amino acids at their C-terminus are recognized as ER residents and are returned to the ER. Proteins destined for mitochondria have signal sequences of yet another type, in which positively charged amino acids alternate with hydrophobic ones. Finally, many proteins destined for peroxisomes have a signal sequence of three characteristic amino acids at their C-terminus.

Table 12–3 presents some specific signal sequences. Experiments in which the peptide is transferred from one protein to another by genetic engineering techniques have demonstrated the importance of each of these signal sequences for protein targeting. Placing the N-terminal ER signal sequence at the beginning of a cytosolic protein, for example, redirects the protein to the ER. Signal sequences are therefore both necessary and sufficient for protein targeting. Even though their amino acid sequences can vary greatly, the signal sequences of all proteins having the same destination are functionally interchangeable; physical properties, such as hydrophobicity, often seem to be more important in the signal-recognition process than the exact amino acid sequence.

Signal sequences are recognized by complementary sorting receptors that guide proteins to their appropriate destination, where the receptors unload their cargo. The receptors function catalytically: after completing one round of targeting, they return to their point of origin to be reused. Most sorting receptors recognize classes of proteins rather than an individual protein species. They can therefore be viewed as public transportation systems, dedicated to delivering different components to their correct location in the cell.

Panel 12–1 describes the main ways of studying how proteins are directed from the cytosol to a specific compartment and the mechanism of their translocation across membranes.

Most Organelles Cannot Be Constructed *De Novo:* They Require Information in the Organelle Itself

When a cell reproduces by division, it has to duplicate its organelles. In general, cells do this by incorporating new molecules into the existing organelles, thereby enlarging them; the enlarged organelles then divide and are distributed to the two daughter cells. Thus, each daughter cell inherits a complete set of specialized cell membranes from its mother. This inheritance is essential because a cell could not make such membranes from scratch. If the ER were completely removed from a cell, for example, how could the cell reconstruct it? As we discuss later, the membrane proteins that define the ER and perform many of its functions are themselves products of the ER. A new ER could not be made without an existing ER or, at least, a membrane that specifically contains the protein

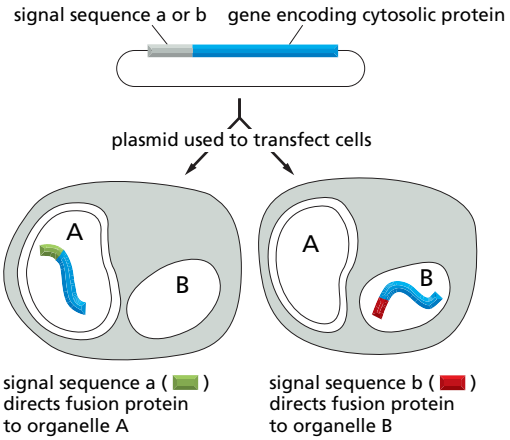
Table 12–3 Some Typical Signal Sequences

FUNCTION OF SIGNAL SEQUENCE	EXAMPLE OF SIGNAL SEQUENCE
Import into nucleus	-Pro-Pro-Lys-Lys-Lys-Arg-Lys-Val-
Export from nucleus	-Leu-Ala-Leu-Lys-Leu-Ala-Gly-Leu-Asp-Ile-
Import into mitochondria	$^{+}H_3N$-Met-Leu-Ser-Leu-Arg-Gln-Ser-Ile-Arg-Phe-Phe-Lys-Pro-Ala-Thr-Arg-Thr-Leu-Cys-Ser-Ser-Arg-Tyr-Leu-Leu-
Import into plastid	$^{+}H_3N$-Met-Val-Ala-Met-Ala-Met-Ala-Ser-Leu-Gln-Ser-Ser-Met-Ser-Ser-Leu-Ser-Leu-Ser-Ser-Asn-Ser-Phe-Leu-Gly-Gln-Pro-Leu-Ser-Pro-Ile-Thr-Leu-Ser-Pro-Phe-Leu-Gln-Gly-
Import into peroxisomes	-Ser-Lys-Leu-COO^{-}
Import into ER	$^{+}H_3N$-Met-Met-Ser-Phe-Val-Ser-Leu-Leu-Leu-Val-Gly-Ile-Leu-Phe-Trp-Ala-Thr-Glu-Ala-Glu-Gln-Leu-Thr-Lys-Cys-Glu-Val-Phe-Gln-
Return to ER	-Lys-Asp-Glu-Leu-COO^{-}

Some characteristic features of the different classes of signal sequences are highlighted in color. Where they are known to be important for the function of the signal sequence, positively charged amino acids are shown in *red* and negatively charged amino acids are shown in *green*. Similarly, important hydrophobic amino acids are shown in *white* and important hydroxylated amino acids are shown in *blue*. $^{+}H_3N$ indicates the N-terminus of a protein; COO^{-} indicates the C-terminus.

A TRANSFECTION APPROACH FOR DEFINING SIGNAL SEQUENCES

One way to show that a signal sequence is required and sufficient to target a protein to a specific intracellular compartment is to create a fusion protein in which the signal sequence is attached by genetic engineering techniques to a protein that normally resides in the cytosol. After the cDNA encoding this protein is transfected into cells, the location of the fusion protein is determined by immunostaining or by cell fractionation.



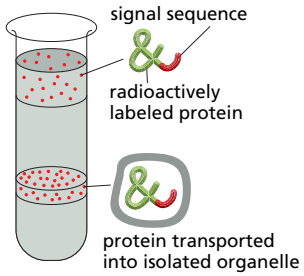
By altering the signal sequence using site-directed mutagenesis, we can determine which structural features are important for its function.

A BIOCHEMICAL APPROACH FOR STUDYING THE MECHANISM OF PROTEIN TRANSLOCATION

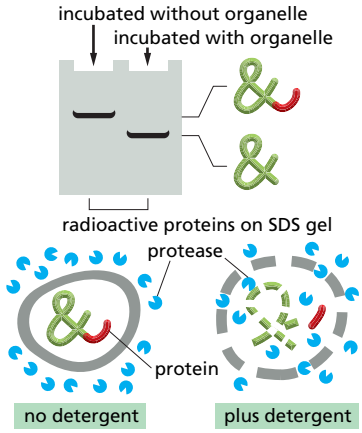
In this approach, a labeled protein containing a specific signal sequence is transported into isolated organelles *in vitro*. The labeled protein is usually produced by cell-free translation of a purified mRNA encoding the protein; radioactive amino acids are used to label the newly synthesized protein so that it can be distinguished from the many other proteins that are present in the *in vitro* translation system.

Three methods are commonly used to test if the labeled protein has been translocated into the organelle:

1. The labeled protein co-fractionates with the organelle during centrifugation.



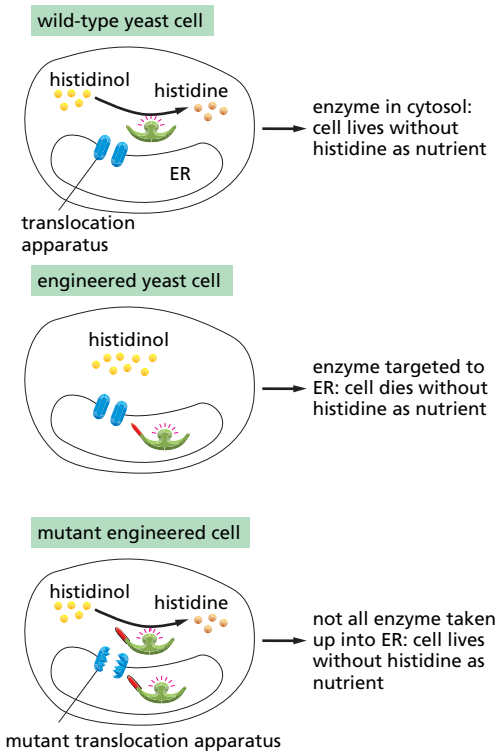
2. The signal sequence is removed by a specific protease that is present inside the organelle.



3. The protein is protected from digestion when proteases are added to the incubation medium, but is susceptible if a detergent is first added to disrupt the organelle membrane.

By exploiting such *in vitro* assays, one can determine what components (proteins, ATP, GTP, etc.) are required for the translocation process.

GENETIC APPROACHES FOR STUDYING THE MECHANISM OF PROTEIN TRANSLOCATION



Yeast cells with mutations in genes that encode components of the translocation machinery have been useful for studying protein translocation. Because mutant cells that cannot translocate any proteins across their membranes will die, the challenge is to find mutations that cause only a partial defect in protein translocation.

One experimental strategy uses genetic engineering to design special yeast cells. The enzyme histidinol dehydrogenase, for example, normally resides in the cytosol, where it is required to produce the essential amino acid histidine from its precursor histidinol. A yeast strain is constructed in which the histidinol dehydrogenase gene is replaced by a re-engineered gene encoding a fusion protein with an added signal sequence that misdirects the enzyme into the endoplasmic reticulum (ER). When such cells are grown without histidine, they die because all of the histidinol dehydrogenase is sequestered in the ER, where it is of no use. Cells with a mutation that only partially inactivates the mechanism for translocating proteins from the cytosol to the ER, however, will survive because the cytosol retains enough of the dehydrogenase to produce histidine.

Often one obtains a cell in which the mutant protein in the translocation machinery still functions partially at normal temperature but is completely inactive at higher temperature. A cell carrying such a temperature-sensitive mutation dies at higher temperature, whether or not histidine is present, as it cannot transport any protein into the ER. The normal gene that was disabled by the temperature-sensitive mutation can be identified by transfecting the mutant cells with a yeast plasmid vector into which random yeast genomic DNA fragments have been cloned: the specific DNA fragment that rescues the mutant cells when they are grown at high temperature should encode the wild-type version of the mutant gene.

translocators required to import selected proteins into the ER from the cytosol (including the ER-specific translocators themselves). The same is true for mitochondria and plastids (see Figure 12–6).

Thus, it seems that the information required to construct an organelle does not reside exclusively in the DNA that specifies the organelle's proteins. Information in the form of at least one distinct protein that preexists in the organelle membrane is also required, and this information is passed from parent cell to progeny cell in the form of the organelle itself. Presumably, such information is essential for the propagation of the cell's compartmental organization, just as the information in DNA is essential for the propagation of the cell's nucleotide and amino acid sequences.

As we discuss in more detail in Chapter 13, however, the ER sheds a constant stream of membrane vesicles that incorporate only a subset of ER proteins and therefore have a different composition from the ER itself. Similarly, the plasma membrane constantly produces various types of specialized endocytic vesicles. Thus, some organelles can form from other organelles and do not have to be inherited at cell division.

Summary

Eucaryotic cells contain intracellular membranes that enclose nearly half the cell's total volume in separate intracellular compartments called organelles. The main types of organelles present in all eucaryotic cells are the endoplasmic reticulum, Golgi apparatus, nucleus, mitochondria, lysosomes, endosomes, and peroxisomes; plant cells also contain plastids, such as chloroplasts. Each organelle contains a distinct set of proteins, which mediate its unique functions.

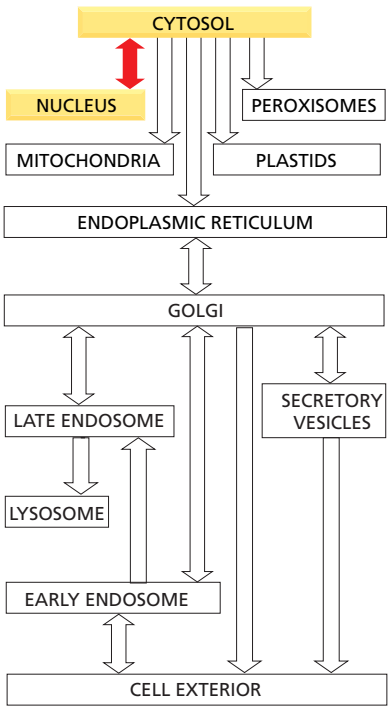
Each newly synthesized organelle protein must find its way from a ribosome in the cytosol, where the protein is made, to the organelle where it functions. It does so by following a specific pathway, guided by sorting signals in its amino acid sequence that function as signal sequences or signal patches. Sorting signals are recognized by complementary sorting receptors, which deliver the protein to the appropriate target organelle. Proteins that function in the cytosol do not contain sorting signals and therefore remain there after they are synthesized.

During cell division, organelles such as the ER and mitochondria are distributed intact to each daughter cell. These organelles contain information that is required for their construction, and so they cannot be made de novo.

THE TRANSPORT OF MOLECULES BETWEEN THE NUCLEUS AND THE CYTOSOL

The **nuclear envelope** encloses the DNA and defines the *nuclear compartment.* This envelope consists of two concentric membranes, which are penetrated by nuclear pore complexes (**Figure 12–8**). Although the inner and outer nuclear membranes are continuous, they maintain distinct protein compositions. The **inner nuclear membrane** contains specific proteins that act as anchoring sites for chromatin and for the *nuclear lamina,* a protein meshwork that provides structural support for the nuclear envelope. The inner membrane is surrounded by the **outer nuclear membrane**, which is continuous with the membrane of the ER. Like the membrane of the ER (discussed later), the outer nuclear membrane is studded with ribosomes engaged in protein synthesis. The proteins made on these ribosomes are transported into the space between the inner and outer nuclear membranes (the *perinuclear space),* which is continuous with the ER lumen.

Bidirectional traffic occurs continuously between the cytosol and the nucleus. The many proteins that function in the nucleus—including histones, DNA and RNA polymerases, gene regulatory proteins, and RNA-processing proteins—are selectively imported into the nuclear compartment from the cytosol, where they are made. At the same time, tRNAs and mRNAs are synthesized in the



nuclear compartment and then exported to the cytosol. Like the import process, the export process is selective; mRNAs, for example, are exported only after they have been properly modified by RNA-processing reactions in the nucleus. In some cases, the transport process is complex. Ribosomal proteins, for instance, are made in the cytosol and imported into the nucleus, where they assemble with newly made ribosomal RNA into particles. The particles are then exported to the cytosol, where they assemble into ribosomes. Each of these steps requires selective transport across the nuclear envelope.

Nuclear Pore Complexes Perforate the Nuclear Envelope

Large, elaborate structures known as **nuclear pore complexes (NPCs)** perforate the nuclear envelope of all eucaryotes. In animal cells, each NPC has an estimated molecular mass of about 125 million daltons and is composed of about 30 different **NPC proteins**, or *nucleoporins*, which are present in multiple copies and are arranged with a striking octagonal symmetry (**Figure 12–9**).

The nuclear envelope of a typical mammalian cell contains 3000–4000 NPCs, and the total traffic that passes through each NPC is enormous: each NPC can transport up to 500 macromolecules per second and can transport in both directions at the same time. How it coordinates the bidirectional flow of macromolecules to avoid congestion and head-on collisions is not known.

Each NPC contains one or more aqueous passages, through which small water-soluble molecules can diffuse passively. Researchers have determined the effective size of these passages by injecting labeled water-soluble molecules of different sizes into the cytosol and then measuring their rate of diffusion into the nucleus. Small molecules (5000 daltons or less) diffuse in so fast that we can consider the nuclear envelope freely permeable to them. Large proteins, however, traverse the NPC much more slowly; the larger a protein, the more slowly it passes through the NPC. Proteins larger than 60,000 daltons can barely enter by passive diffusion. This size cut-off to free diffusion is thought to result from the NPC structure (**Figure 12–10**). Many NPC proteins that line the central pore contain extensive unstructured regions, which are thought to form a disordered tangle (much like a kelp bed in the ocean), blocking the central opening in the NPC to the passage of large macromolecules, but leaving small openings to allow the diffusion of smaller molecules.

Because many cell proteins are too large to diffuse passively through the NPCs, the nuclear compartment and the cytosol can maintain different complements of proteins. Mature cytosolic ribosomes, for example, are about 30 nm in diameter and thus cannot diffuse through the NPC, confining protein synthesis to the cytosol. But how does the nucleus export newly made ribosomal subunits or import large molecules, such as DNA and RNA polymerases, which have subunit molecular weights of 100,000–200,000 daltons? As we discuss next, these and many other protein and RNA molecules bind to specific receptor proteins that ferry large molecules actively through NPCs.

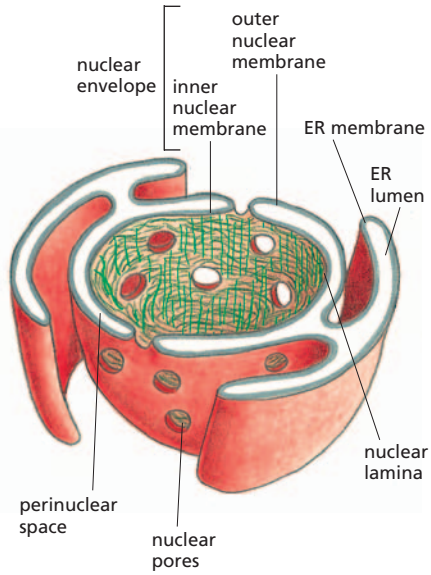


Figure 12–8 The nuclear envelope. The double-membrane envelope is penetrated by pores in which NPCs are positioned and is continuous with the endoplasmic reticulum. The ribosomes that are normally bound to the cytosolic surface of the ER membrane and outer nuclear membrane are not shown. The nuclear lamina is a fibrous meshwork underlying the inner membrane.

Nuclear Localization Signals Direct Nuclear Proteins to the Nucleus

When proteins are experimentally extracted from the nucleus and reintroduced into the cytosol, even the very large ones reaccumulate efficiently in the nucleus. Sorting signals called **nuclear localization signals** are responsible for the selectivity of this active nuclear import process. The signals have been precisely defined by using recombinant DNA technology for numerous nuclear proteins, as well as for proteins that enter the nucleus only transiently (**Figure 12–11**). In many proteins, signals consist of one or two short sequences that are rich in the positively charged amino acids lysine and arginine (see Table 12–3, p. 702), with the precise sequence varying for different nuclear proteins. Other nuclear proteins contain different signals, some of which are not yet characterized.

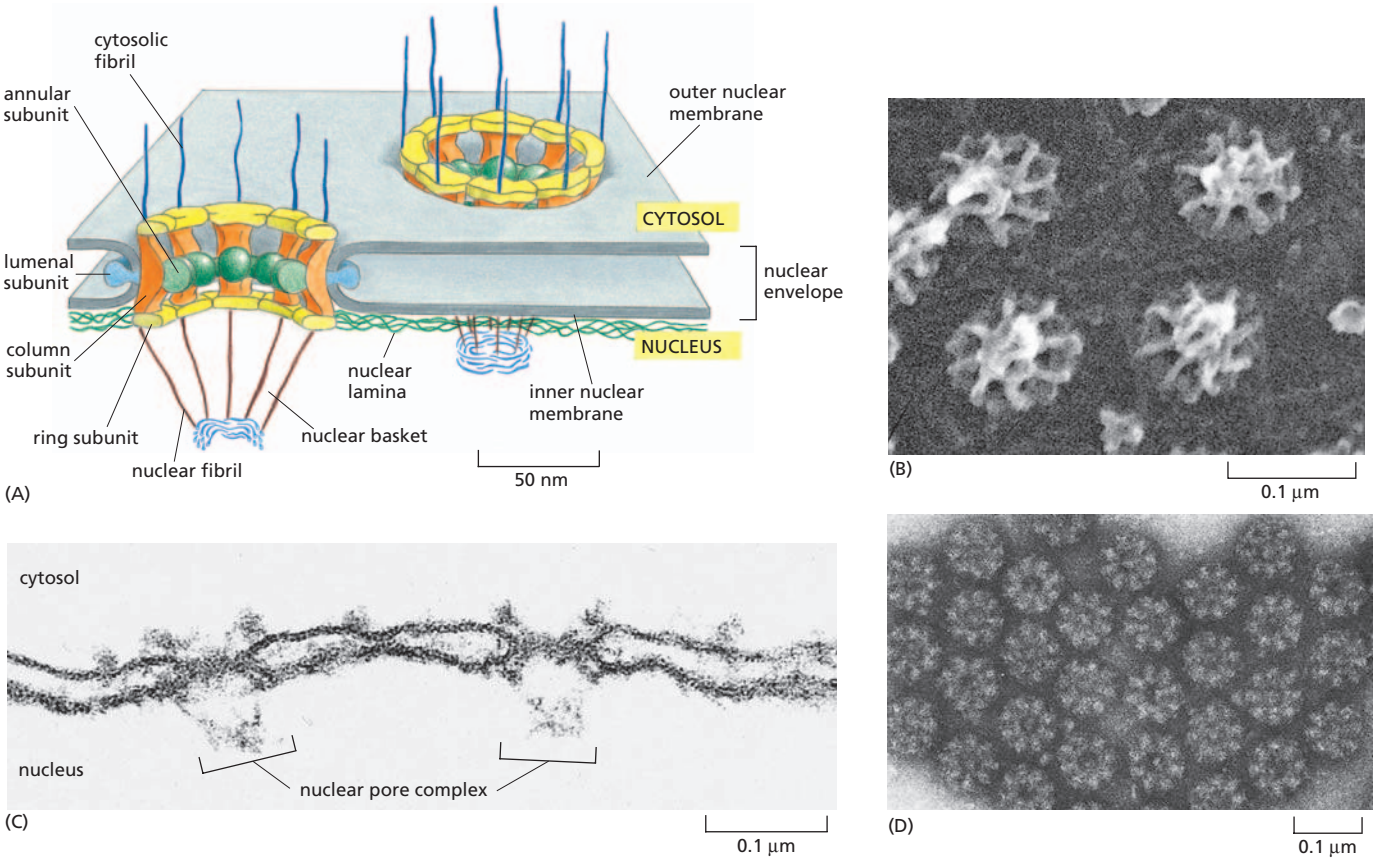


Figure 12–9 The arrangement of NPCs in the nuclear envelope. (A) A small region of the nuclear envelope. In cross section, an NPC seems to have four structural building blocks: column subunits, which form the bulk of the pore wall; annular subunits, which are centrally located; lumenal subunits, which contain transmembrane proteins that anchor the complex to the nuclear membrane; and ring subunits, which form the cytosolic and nuclear faces of the complex. In addition, fibrils protrude from both the cytosolic and the nuclear sides of the NPC. On the nuclear side, the fibrils converge to form basketlike structures. Immunoelectron microscopic studies show that the proteins that make up the core of the NPC are oriented symmetrically across the nuclear envelope, so that the nuclear and cytosolic sides of the core look identical. In contrast, the proteins that make up the fibrils are different on the cytosolic and nuclear sides of the NPC. The eight-fold rotational and two-fold transverse symmetry of the core NPC explains how such a huge structure can be formed from only about 30 different proteins: many of these proteins are present in 16 copies (or multiples of 16). Disordered domains of the core proteins (not shown) are thought to extend toward the center of the NPC, blocking the passive diffusion of large macromolecules. (B) A scanning electron micrograph of the nuclear side of the nuclear envelope of an oocyte (see also Figure 9–55). (C) An electron micrograph showing a side view of two NPCs *(brackets);* note that the inner and outer nuclear membranes are continuous at the edges of the pore. (D) An electron micrograph showing face-on views of negatively stained NPCs. The membrane has been removed by detergent extraction. Note that some of the NPCs contain material in their center, which is thought to be macromolecules in transit through these NPCs. (B, from M.W. Goldberg and T.D. Allen, *J. Cell Biol.* 119:1429–1440, 1992. With permission from The Rockefeller University Press; C, courtesy of Werner Franke and Ulrich Scheer; D, courtesy of Ron Milligan.)

Nuclear localization signals can be located almost anywhere in the amino acid sequence and are thought to form loops or patches on the protein surface. Many function even when linked as short peptides to lysine side chains on the surface of a cytosolic protein, suggesting that the precise location of the signal within the amino acid sequence of a nuclear protein is not important. Moreover, as long as one of the protein subunits of a multicomponent complex displays a nuclear localization signal, the complex can be imported into the nucleus.

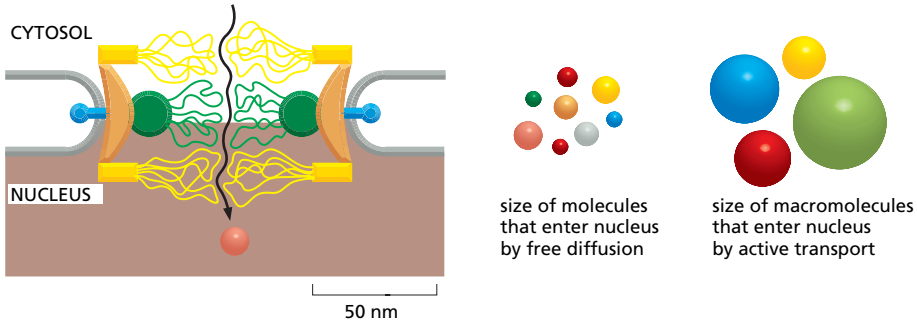


Figure 12–10 A model for the gated diffusion barrier of the NPC. The drawing shows a side view of an NPC. Unstructured regions of the proteins lining the central pore form a tangled meshwork, which blocks the passive diffusion of large macromolecules. During active transport, however, even particles up to 39 nm in diameter can pass through NPCs.

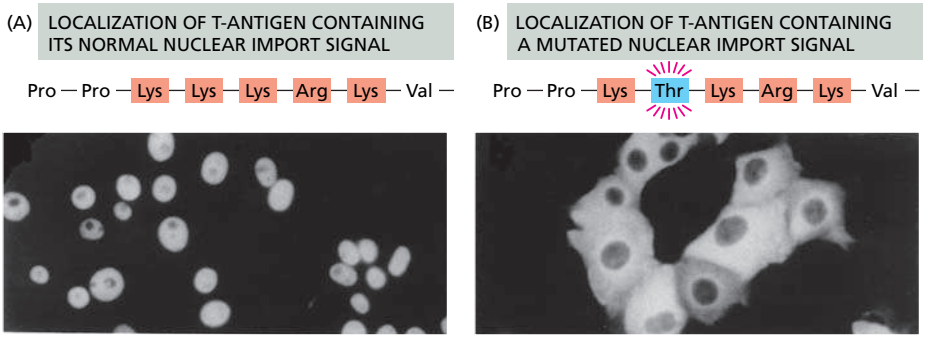


Figure 12–11 The function of a nuclear localization signal. Immunofluorescence micrographs showing the cell location of SV40 virus T-antigen containing or lacking a short sequence that serves as a nuclear localization signal. (A) The normal T-antigen protein contains the lysine-rich sequence indicated and is imported to its site of action in the nucleus, as indicated by immunofluorescence staining with antibodies against the T-antigen. (B) T-antigen with an altered nuclear localization signal (a threonine replacing a lysine) remains in the cytosol. (From D. Kalderon, B. Roberts, W. Richardson and A. Smith, *Cell* 39:499–509, 1984. With permission from Elsevier.)

One can make the transport of nuclear proteins through NPCs visible by coating gold particles with a nuclear localization signal, injecting the particles into the cytosol, and then following their fate by electron microscopy (**Figure 12–12**). Transport begins when the particles bind to tentaclelike fibrils that extend from the rim of the NPC into the cytoplasm, and then proceed through the center of the NPC. Presumably, the unstructured regions of the NPC proteins that form a diffusion barrier for large molecules (mentioned earlier) are pushed away to allow the coated gold particles to squeeze through.

Macromolecular transport across NPCs differs fundamentally from the transport of proteins across the membranes of other organelles, in that it occurs through a large aqueous pore rather than through a protein transporter spanning one or more lipid bilayers. For this reason, fully folded nuclear proteins can be transported into the nucleus through an NPC, and newly formed ribosomal subunits are transported out of the nucleus as an assembled particle. By contrast, proteins have to be extensively unfolded to be transported into most other organelles, as we discuss later. In the electron microscope, however, very large particles traversing the NPC seem to become compressed as they squeeze through the pore, indicating that they undergo restructuring during transport. The export of some very large mRNAs has been extensively studied, as discussed in Chapter 6 (see Figure 6–39).

Nuclear Import Receptors Bind to Both Nuclear Localization Signals and NPC proteins

To initiate nuclear import, most nuclear localization signals must be recognized by **nuclear import receptors**, which are encoded by a family of related genes. Each family member encodes a receptor protein that is specialized for the transport of a subset of cargo proteins—those that contain nuclear localization signals to which the receptor can bind (**Figure 12–13**A).

The import receptors are soluble cytosolic proteins that bind both to the nuclear localization signal on the protein to be transported and to NPC proteins, some of which form the fibrils of the NPC that extend into the cytosol. These fibrils, as well as many of the NPC proteins that line the center of the NPC and contribute to the diffusion barrier, include a large number of short amino-acid repeats that contain phenylalanine and glycine and are therefore called *FG-repeats* (named after the one-letter code for these amino acids, discussed in Chapter 3). FG-repeats serve as binding sites for the import receptors. They are thought to line the path through the NPCs taken by the import receptors and their bound cargo proteins. The receptor–cargo complexes move along the transport path by repeatedly binding, dissociating, and then re-binding to adjacent FG-repeat sequences. In this way, the complexes hop from one NPC protein to another to traverse the tangled interior of the NPC. Once inside the nucleus, the import receptors dissociate from their cargo and return to the cytosol.

Nuclear import receptors do not always bind to nuclear proteins directly. Additional adaptor proteins sometimes form a bridge between the import receptors and the nuclear localization signals on the proteins to be transported (Figure 12–13B). Some adaptor proteins are structurally related to nuclear import receptors, suggesting a common evolutionary origin. By using a variety of different

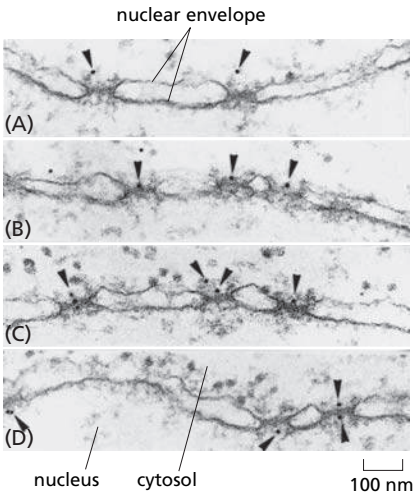


Figure 12–12 Visualizing active import through NPCs. This series of electron micrographs shows colloidal gold spheres *(arrowheads)* coated with peptides containing nuclear localization signals entering the nucleus through NPCs. Gold particles were injected into living cells, which then were fixed and prepared for electron microscopy at various times after injection. (A) Gold particles are first seen in proximity to the cytosolic fibrils of the NPCs. (B, C) They migrate to the center of the NPCs, where they are first seen exclusively on the cytosolic face. (D) They then appear on the nuclear face. These gold particles have much larger diameters than the diffusion channel in the NPC and are imported by active transport. (From N. Panté and U. Aebi, *Science* 273:1729–1732, 1996. With permission from AAAS.)

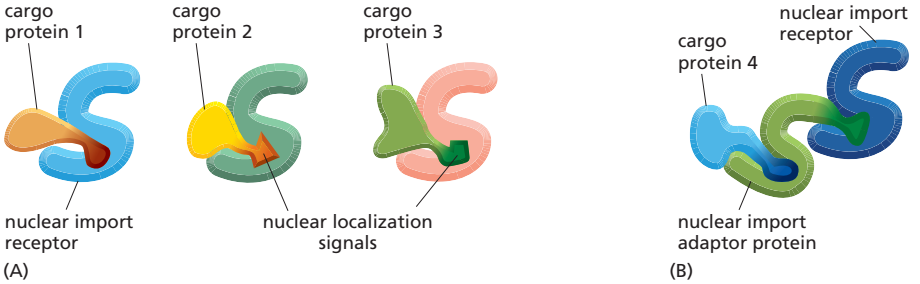


Figure 12–13 Nuclear import receptors. (A) Many nuclear import receptors bind both to NPC proteins and to a nuclear localization signal on the cargo proteins they transport. Cargo proteins 1, 2, and 3 in this example contain different nuclear localization signals, and therefore each binds to different nuclear import receptors. (B) Cargo protein 4 requires an adaptor protein to bind to its nuclear import receptor. The adaptors are structurally related to nuclear import receptors and recognize nuclear localization signals on cargo proteins. They also contain a nuclear localization signal that binds them to an import receptor.

import receptors and adaptors, cells are able to recognize the broad repertoire of nuclear localization signals that are displayed on nuclear proteins.

Nuclear Export Works Like Nuclear Import, But in Reverse

The nuclear export of large molecules, such as new ribosomal subunits and RNA molecules, occurs through NPCs and also depends on a selective transport system. The transport system relies on **nuclear export signals** on the macromolecules to be exported, as well as on complementary **nuclear export receptors**. These receptors bind to both the export signal and NPC proteins to guide their cargo through the NPC to the cytosol.

Many nuclear export receptors are structurally related to nuclear import receptors, and they are encoded by the same gene family of **nuclear transport receptors**, or *karyopherins*. In yeast, there are 14 genes encoding members of this family; in animal cells, the number is significantly larger. It is often not possible to tell from their amino acid sequence alone whether a particular family member works as a nuclear import or nuclear export receptor. As might be expected, therefore, the import and export transport systems work in similar ways but in opposite directions: the import receptors bind their cargo molecules in the cytosol, release them in the nucleus, and are then exported to the cytosol for reuse, while the export receptors function in the opposite fashion.

The Ran GTPase Imposes Directionality on Transport Through NPCs

The import of nuclear proteins through the NPC concentrates specific proteins in the nucleus and thereby increases order in the cell. The cell obtains the energy it needs for this process through the hydrolysis of GTP by the monomeric GTPase **Ran**. Ran is found in both the cytosol and the nucleus, and it is required for both nuclear import and export.

Like other GTPases, Ran is a molecular switch that can exist in two conformational states, depending on whether GDP or GTP is bound (discussed in Chapter 3). Two Ran-specific regulatory proteins trigger the conversion between the two states: a cytosolic *GTPase-activating protein (GAP)* triggers GTP hydrolysis and thus converts Ran-GTP to Ran-GDP, and a nuclear *guanine exchange factor (GEF)* promotes the exchange of GDP for GTP and thus converts Ran-GDP to Ran-GTP. Because *Ran-GAP* is located in the cytosol and *Ran-GEF* is located in the nucleus, the cytosol contains mainly Ran-GDP, and the nucleus contains mainly Ran-GTP (**Figure 12–14**).

This gradient of the two conformational forms of Ran drives nuclear transport in the appropriate direction (**Figure 12–15**). Docking of nuclear import receptors to FG-repeats on the cytosolic side of the NPC, for example, occurs whether or not these receptors are loaded with an appropriate cargo. Import receptors then hop from FG-repeat to FG-repeat. If they reach the nuclear side of the pore complex, Ran-GTP binds to them, and, if they arrive loaded with cargo molecules, the Ran-GTP binding causes the import receptors to release their cargo (**Figure 12–16**). Because the Ran-GDP in the cytosol does not bind to the cargo receptors, unloading occurs only on the nuclear side of the NPC. In this way, the nuclear localization of Ran-GTP creates the directionality.

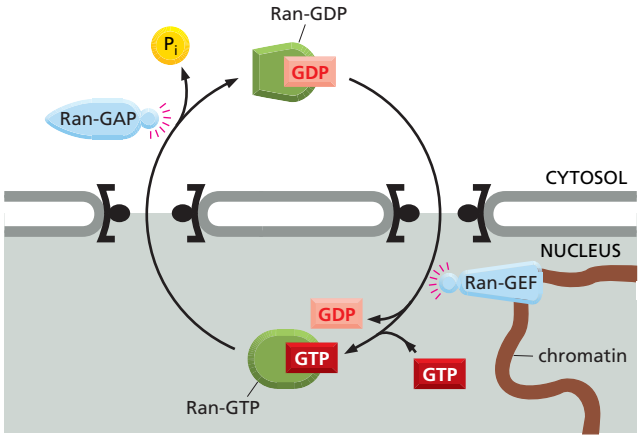


Figure 12–14 The compartmentalization of Ran-GDP and Ran-GTP. Localization of Ran-GDP in the cytosol and Ran-GTP in the nucleus results from the localization of two Ran regulatory proteins: Ran GTPase-activating protein (Ran-GAP) is located in the cytosol and Ran guanine nucleotide exchange factor (Ran-GEF) binds to chromatin and is therefore located in the nucleus.

Having discharged its cargo in the nucleus, the empty import receptor with Ran-GTP bound is transported back through the pore complex to the cytosol. There, Ran-GAP triggers Ran-GTP to hydrolyze its bound GTP, thereby converting it to Ran-GDP. The import receptor is then ready for another cycle of nuclear import.

Nuclear export occurs by a similar mechanism, except that Ran-GTP in the nucleus promotes cargo binding to the export receptor, rather than dissociating it. Once the export receptor moves through the pore to the cytosol, its Ran-GTP encounters Ran-GAP and hydrolyzes GTP. As a result, the export receptor releases both its cargo and Ran-GDP in the cytosol. Free export receptors are then returned to the nucleus to complete the cycle (see Figure 12–15).

Transport Through NPCs Can Be Regulated by Controlling Access to the Transport Machinery

Some proteins, such as those that bind newly made mRNAs in the nucleus, contain both nuclear localization signals and nuclear export signals. These proteins continually shuttle back and forth between the nucleus and the cytosol. The relative rates of their import and export determine the steady-state localization of such *shuttling proteins*. If the rate of import exceeds the rate of export, a protein will be located mainly in the nucleus. Conversely, if the rate of export exceeds the rate of import, a protein will be located mainly in the cytosol. Thus, changing the rate of import, export, or both, can change the location of a protein.

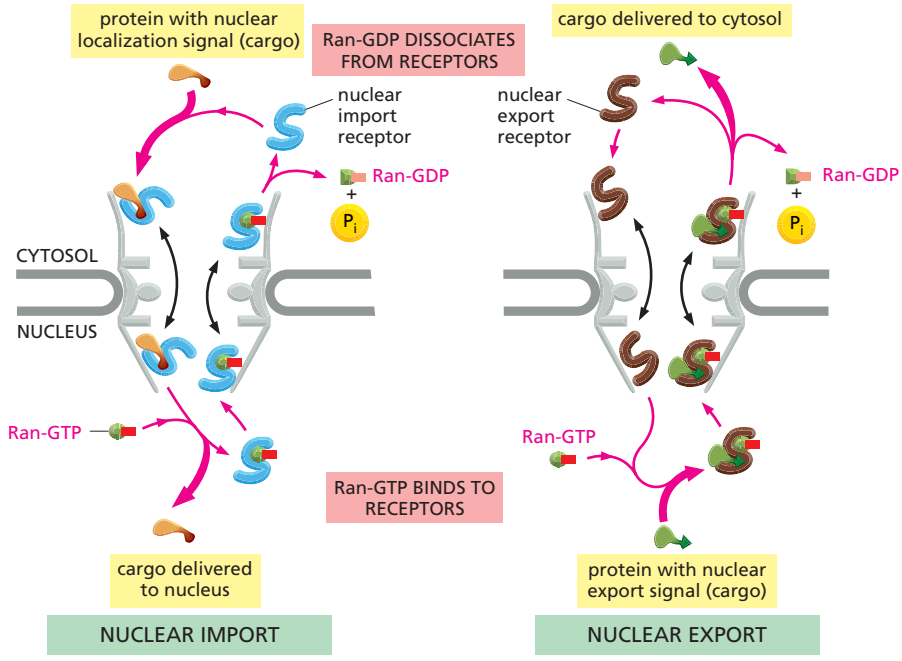


Figure 12–15 A model explaining how GTP hydrolysis by Ran in the cytosol provides directionality to nuclear transport. Movement through the NPC of loaded nuclear transport receptors occurs along the FG-repeats displayed by certain NPC proteins. The differential localization of Ran-GTP in the nucleus and Ran-GDP in the cytosol provides directionality *(red arrows)* to both nuclear import *(left)* and nuclear export *(right)*. Ran-GAP stimulates the hydrolysis of GTP to produce Ran-GDP on the cytosolic side of the NPC (see Figure 12–14).

Ran-GDP is imported into the nucleus by its own import receptor, which is specific for the GDP-bound conformation of Ran. The Ran-GDP receptor is structurally unrelated to the main family of nuclear transport receptors. However, it also binds to FG-repeats in NPC proteins and hops across the NPC.

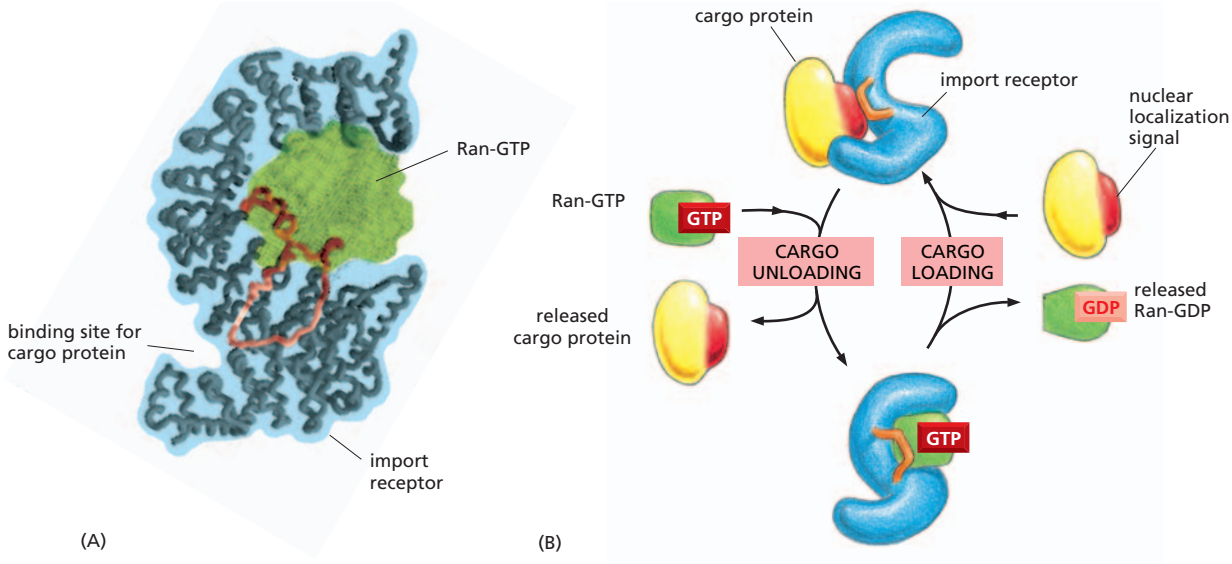


Figure 12–16 How the binding of Ran-GTP can cause nuclear import receptors to release their cargo. (A) The structure of a nuclear transport receptor with bound Ran-GTP. The receptor contains repeated α-helical motifs that stack on top of each other, forming a flexible springlike structure, which can adopt multiple conformations in response to the binding of cargo proteins and Ran-GTP. Note that cargo proteins and Ran-GTP bind to different regions of the coiled spring. The Ran-GTP partly covers a conserved loop *(red)* of the receptor, which, in the Ran-free state, is thought to be important for cargo binding. (B) The cycle of loading in the cytosol and unloading in the nucleus of a nuclear import receptor. (A, adapted from Y.M. Chook and G. Blobel, *Nature* 399:230–237, 1999. With permission from Macmillan Publishers Ltd.)

Some shuttling proteins move continuously into and out of the nucleus. In other cases, however, the transport is stringently controlled. As discussed in Chapter 7, cells control the activity of some gene regulatory proteins by keeping them out of the nucleus until they are needed there (**Figure 12–17**). In many cases, cells control transport by regulating nuclear localization and export signals; these can be turned on or off, often by phosphorylation of amino acids close to the signal sequences (**Figure 12–18**).

Other gene regulatory proteins are bound to inhibitory cytosolic proteins that either anchor them in the cytosol (through interactions with the cytoskeleton or specific organelles) or mask their nuclear localization signals so that they cannot interact with nuclear import receptors. An appropriate stimulus releases the gene regulatory protein from its cytosolic anchor or mask, and it is then transported into the nucleus. One important example is the latent gene regulatory protein that controls the expression of proteins involved in cholesterol metabolism. The protein is made and stored in an inactive form as a transmembrane protein in the ER. When deprived of cholesterol, the protein exits the ER to the Golgi apparatus where it meets specific proteases that cleave it, releasing its cytosolic domain into the cytosol. This domain is then imported into the nucleus, where it activates the transcription of genes required for both cholesterol import and synthesis.

As we discussed in detail in Chapter 6, cells control the export of mRNA from the nucleus in a similar way. Proteins that guide the mRNA out of the nucleus load onto the RNA as transcription and splicing proceed. Upon entry into the cytosol, the proteins are stripped off and rapidly returned to the nucleus. Other RNAs, such as snRNAs and tRNAs, are exported by different nuclear export receptors.

During Mitosis the Nuclear Envelope Disassembles

The **nuclear lamina**, located on the nuclear side of the inner nuclear membrane, is a meshwork of interconnected protein subunits called **nuclear lamins**. The lamins are a special class of intermediate filament proteins (discussed in Chapter 16) that polymerize into a two-dimensional lattice (**Figure 12–19**). The nuclear lamina gives shape and stability to the nuclear envelope, to which it is anchored by attachment to both the NPCs and integral membrane proteins of the inner nuclear membrane. The lamina also interacts directly with chromatin, which

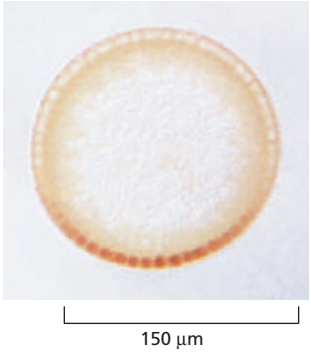


Figure 12–17 The control of nuclear transport in the fly embryo. The gene regulatory protein *dorsal* is expressed uniformly throughout this early *Drosophila* embryo, which is shown in cross section. The protein has been stained with an enzyme-coupled antibody that yields a brown product. It is active only in cells at the ventral side *(bottom)* of the embryo, where it is found in nuclei. (Courtesy of Siegfried Roth.)

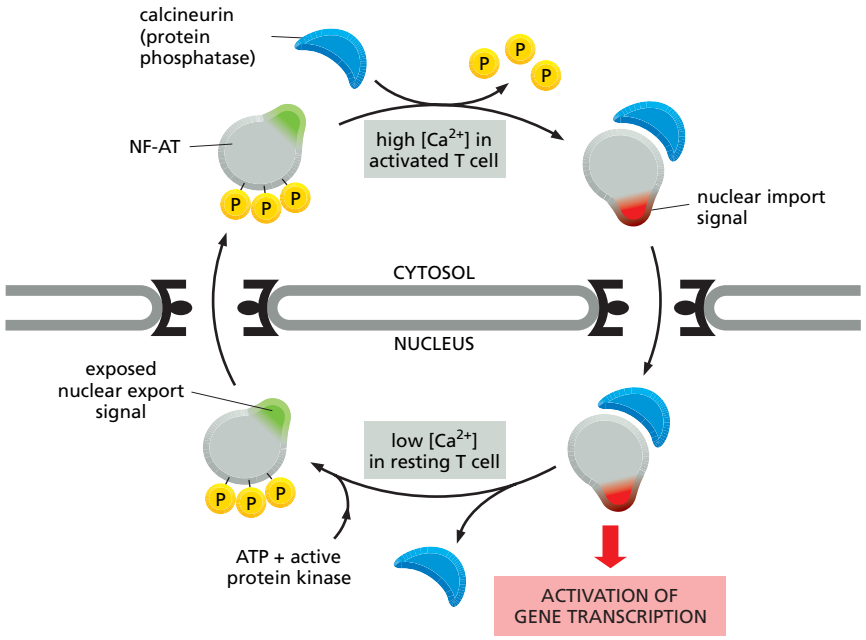


Figure 12–18 The control of nuclear import during T cell activation. <AGTT> The nuclear factor of activated T cells (NF-AT) is a gene regulatory protein that, in the resting T cell, is found in the cytosol in a phosphorylated state. When T cells are activated by foreign antigen (discussed in Chapter 25), the intracellular Ca^{2+} concentration increases. In high Ca^{2+}, the protein phosphatase calcineurin binds to NF-AT and dephosphorylates it. The dephosphorylation exposes nuclear import signals and blocks a nuclear export signal. The complex of NF-AT and calcineurin is then imported into the nucleus, where NF-AT activates the transcription of numerous genes required for T cell activation.

The response shuts off when Ca^{2+} levels decrease, releasing NF-AT from calcineurin. Rephosphorylation of NF-AT inactivates the nuclear import signals and re-exposes the nuclear export signal, causing NF-AT to relocate to the cytosol. Some of the most potent immunosuppressive drugs, including cyclosporin A and FK506, inhibit the ability of calcineurin to dephosphorylate NF-AT and thereby block the nuclear accumulation of NF-AT and T cell activation.

itself interacts with integral membrane proteins of the inner nuclear membrane. Together with the lamina, these inner membrane proteins provide structural links between the DNA and the nuclear envelope.

When a nucleus disassembles during mitosis, the nuclear lamina depolymerizes. The disassembly is at least partly a consequence of direct phosphorylation of the nuclear lamins by the cyclin-dependent protein kinase Cdk that is activated at the onset of mitosis (discussed in Chapter 16). At the same time, proteins of the inner nuclear membrane are phosphorylated, and the NPCs disassemble and disperse in the cytosol. During this process, some NPC proteins become bound to nuclear import receptors, which play an important part in the reassembly of NPCs at the end of mitosis. Nuclear envelope membrane proteins—no longer tethered to the pore complexes, lamina, or chromatin—disperse throughout the ER membrane. The dynein motor protein, which moves along microtubules (discussed in Chapter 17), actively participates in tearing the nuclear envelope off the chromatin. Together, these processes break down the barriers that normally separate the nucleus and cytosol, and the nuclear proteins that are not bound to membranes or chromosomes intermix completely with those of the cytosol (**Figure 12–20**).

Later in mitosis, the nuclear envelope reassembles on the surface of the chromosomes. In addition to its crucial role in nuclear transport, Ran-GTPase also acts as a positional marker for chromatin during cell division, when the nuclear and cytosolic components intermix. Because Ran-GEF remains bound to chromatin when the nuclear envelope breaks down, Ran molecules close to chromatin are mainly in their GTP-bound conformation. By contrast, Ran molecules further away have a high likelihood of encountering Ran-GAP, which is distributed throughout the cytosol; these Ran molecules are therefore mainly in their GDP-bound conformation.

Chromatin in mitotic cells is therefore surrounded by a cloud of Ran-GTP. This cloud locally displaces nuclear import receptors from NPC proteins, which starts the assembly process of NPCs attached to the chromosome surface. At the same time, inner nuclear membrane proteins and dephosphorylated lamins bind again to chromatin. ER membranes wrap around groups of chromosomes and continue fusing until they form a sealed nuclear envelope. During this process, the NPCs start actively reimporting proteins that contain nuclear localization signals. Because the nuclear envelope is initially closely applied to the surface of the chromosomes, the newly formed nucleus excludes all proteins except those initially bound to the mitotic chromosomes and those that are selectively imported through NPCs. In this way, all other large proteins are kept out of the newly assembled nucleus.

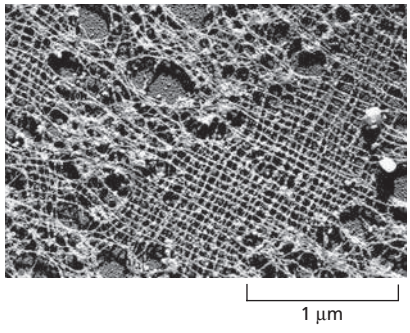


Figure 12–19 The nuclear lamina. An electron micrograph of a portion of the nuclear lamina in a *Xenopus* oocyte prepared by freeze-drying and metal shadowing. The lamina is formed by a regular lattice of specialized intermediate filaments. (Courtesy of Ueli Aebi.)

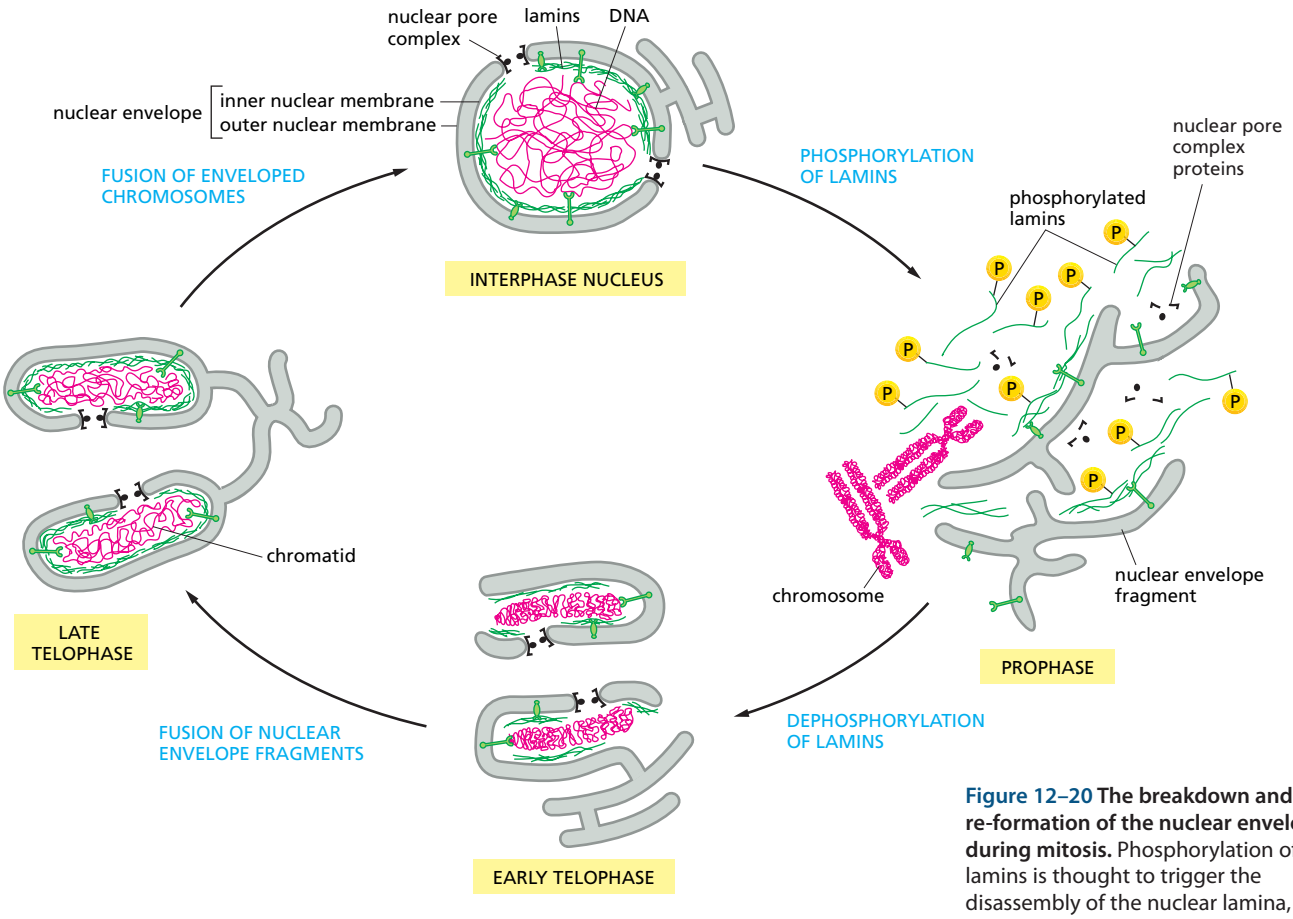


Figure 12–20 The breakdown and re-formation of the nuclear envelope during mitosis. Phosphorylation of the lamins is thought to trigger the disassembly of the nuclear lamina, which helps the nuclear envelope to break up. Dephosphorylation of the lamins is thought to help reverse the process. An analogous phosphorylation and dephosphorylation cycle occurs for some nucleoporins and proteins of the inner nuclear membrane, and some of these dephosphorylations are also involved in the reassembly process shown.

As indicated, the nuclear envelope initially reforms around individual decondensing chromatids. Eventually as decondensation progresses, these structures fuse to form a single complete nucleus.

Nuclear localization signals are not cleaved off after transport into the nucleus, presumably because nuclear proteins need to be imported repeatedly, once after every cell division. In contrast, once a protein molecule has been imported into any of the other membrane-enclosed organelles, it is passed on from generation to generation within that compartment and does not need to be translocated ever again; the signal sequence on these molecules is often removed after protein translocation.

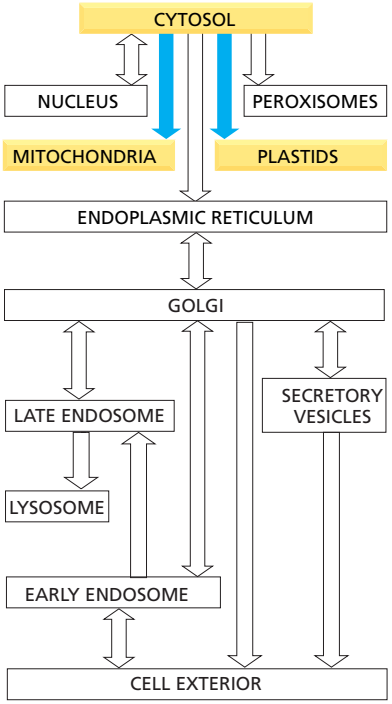
As we discuss in Chapter 17, the cloud of Ran-GTP surrounding chromatin is also important in assembling the mitotic spindle in a dividing cell.

Summary

The nuclear envelope consists of an inner and an outer nuclear membrane. The outer membrane is continuous with the ER membrane, and the space between it and the inner membrane is continuous with the ER lumen. RNA molecules, which are made in the nucleus, and ribosomal subunits, which are assembled there, are exported to the cytosol; in contrast, all the proteins that function in the nucleus are synthesized in the cytosol and are then imported. The extensive traffic of materials between the nucleus and cytosol occurs through nuclear pore complexes (NPCs), which provide a direct passageway across the nuclear envelope. Small molecules diffuse passively through the NPCs, but large macromolecules have to be actively transported.

Proteins containing nuclear localization signals are actively transported inward through NPCs, while RNA molecules and newly made ribosomal subunits contain nuclear export signals, which direct their outward active transport through NPCs. Some proteins, including the nuclear import and export receptors, continually shuttle between the cytosol and nucleus. Ran-GTPase provides both the free energy and the directionality for nuclear transport. Cells regulate the transport of nuclear proteins and RNA molecules through the NPCs by controlling the access of these molecules to the transport machinery. Because nuclear localization signals are not removed, nuclear proteins can be imported repeatedly, as is required each time that the nucleus reassembles after mitosis.

THE TRANSPORT OF PROTEINS INTO MITOCHONDRIA AND CHLOROPLASTS



Mitochondria and chloroplasts are double-membrane-enclosed organelles (discussed in Chapter 14). They specialize in ATP synthesis, using energy derived from electron transport and oxidative phosphorylation in mitochondria and from photosynthesis in chloroplasts. Although both organelles contain their own DNA, ribosomes, and other components required for protein synthesis, most of their proteins are encoded in the cell nucleus and imported from the cytosol. Each imported protein must reach the particular organelle subcompartment in which it functions.

There are two subcompartments in mitochondria: the internal **matrix space** and the **intermembrane space**. These compartments are formed by the two concentric mitochondrial membranes: the **inner membrane**, which encloses the matrix space and forms extensive invaginations called *cristae*, and the **outer membrane**, which is in contact with the cytosol (**Figure 12–21**A). Chloroplasts have the same two subcompartments plus an additional subcompartment, the *thylakoid space*, which is surrounded by the *thylakoid membrane* (Figure 12–21B). Each of the subcompartments in mitochondria and chloroplasts contains a distinct set of proteins.

New mitochondria and chloroplasts are produced by the growth of preexisting organelles, followed by fission (discussed in Chapter 14). The growth depends mainly on the import of proteins from the cytosol. The imported proteins must be transported across a number of membranes in succession and end up in the appropriate place. The process of protein movement across membranes is often called *protein translocation*. This section explains how it occurs.

Translocation into Mitochondria Depends on Signal Sequences and Protein Translocators

Proteins imported into **mitochondria** are usually taken up from the cytosol within seconds or minutes of their release from ribosomes. Thus, in contrast to the protein translocation into the ER, described later, mitochondrial proteins are first fully synthesized as **mitochondrial precursor proteins** in the cytosol and then translocated into mitochondria by a *post-translational* mechanism. One or more signal sequences direct all mitochondrial precursor proteins to their appropriate mitochondrial subcompartment. Many proteins entering the matrix space contain a signal sequence at their N-terminus that a signal peptidase rapidly removes after import. Others, including all outer membrane and many inner membrane and intermembrane space proteins, have an internal signal sequence that is not removed. The signal sequences are both necessary and sufficient for the import and correct localization of the proteins: when genetic engineering techniques are used to link these signals to a cytosolic protein, the signals direct the protein to the correct mitochondrial subcompartment.

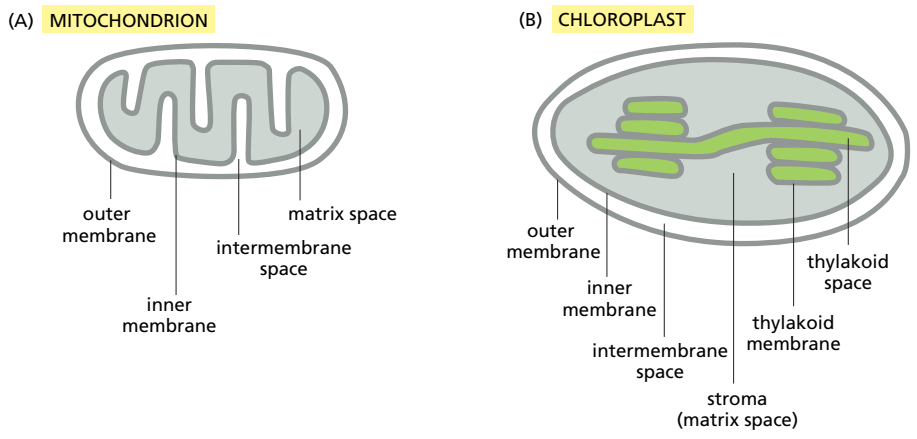


Figure 12–21 The subcompartments of mitochondria and chloroplasts. In contrast to the cristae of mitochondria (A), the thylakoids of chloroplasts (B) are not connected to the inner membrane and therefore form a compartment with a separate internal space (see Figure 12–3).

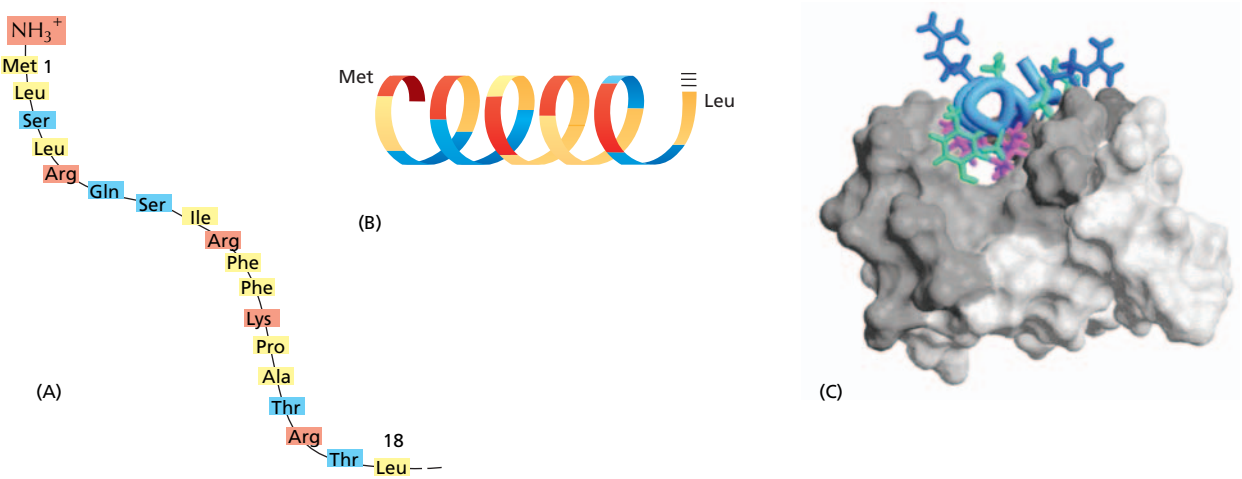


Figure 12–22 A signal sequence for mitochondrial protein import. Cytochrome oxidase is a large multiprotein complex located in the inner mitochondrial membrane, where it functions as the terminal enzyme in the electron-transport chain (discussed in Chapter 14). (A) The first 18 amino acids of the precursor to subunit IV of this enzyme serve as a signal sequence for import of the subunit into the mitochondrion. (B) When the signal sequence is folded as an α helix, the positively charged residues *(red)* are clustered on one face of the helix, while the nonpolar residues *(yellow)* are clustered primarily on the opposite face. Amino acids with uncharged polar side chains are shaded *blue.* Signal sequences that direct proteins into the matrix space always have the potential to form such an amphiphilic α helix, which is recognized by specific receptor proteins on the mitochondrial surface. (C) The structure of a signal sequence of alcohol dehydrogenase, another mitochondrial matrix enzyme, bound to an import receptor was determined by NMR. The amphiphilic α helix binds with its hydrophobic face to a hydrophobic groove in the receptor. (C, adapted from Y. Abe et al., *Cell* 100:551–560, 2000. With permission from Elsevier.)

The signal sequences that direct precursor proteins into the mitochondrial matrix space are best understood. They all form an amphiphilic α helix, in which positively charged residues cluster on one side of the helix, while uncharged hydrophobic residues cluster on the opposite side. Specific receptor proteins that initiate protein translocation recognize this configuration rather than the precise amino acid sequence of the signal sequence (**Figure 12–22**).

Multisubunit protein complexes that function as **protein translocators** mediate protein translocation across mitochondrial membranes. The **TOM complex** transfers proteins across the outer membrane, and two **TIM complexes** (TIM23 and TIM22) transfer proteins across the inner membrane (**Figure 12–23**). (TOM and TIM stand for translocase of the outer and inner mitochondrial membranes, respectively.) These complexes contain some components that act as receptors for mitochondrial precursor proteins, and other components that form the translocation channels.

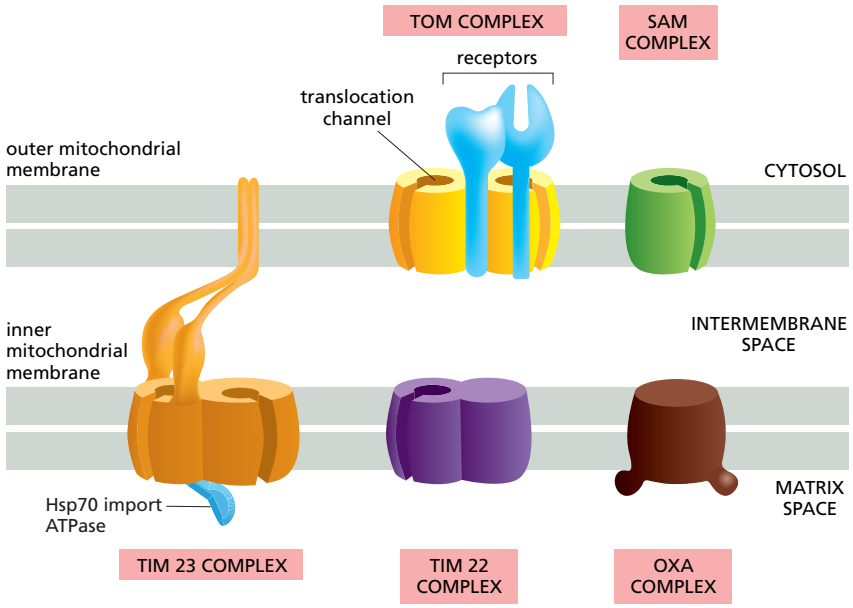


Figure 12–23 The protein translocators in the mitochondrial membranes. The TOM, TIM, SAM, and OXA complexes are multimeric membrane protein assemblies that catalyze protein transport across mitochondrial membranes. The protein components of the TIM22 and TIM23 complexes that line the import channel are structurally related, suggesting a common evolutionary origin of both TIM complexes. As indicated, one of the core components of the TIM23 complex contains a hydrophobic α-helical extension that is inserted into the outer mitochondrial membrane; the complex is therefore unusual in that it simultaneously spans two membranes.

On the matrix side, the TIM23 complex is bound to a protein complex containing mitochondrial Hsp70, which acts as an import ATPase using ATP hydrolysis to pull proteins through the pore.

The TOM complex is required for the import of all nucleus-encoded mitochondrial proteins. It initially transports their signal sequences into the intermembrane space and helps to insert transmembrane proteins into the outer membrane. β-Barrel proteins, which are particularly abundant in the outer membrane, are then passed on to an additional translocator, the **SAM complex**, which helps them to fold properly in the outer membrane. The TIM23 complex transports some soluble proteins into the matrix space and helps to insert transmembrane proteins into the inner membrane. The TIM22 complex mediates the insertion of a subclass of inner membrane proteins, including the transporter that moves ADP, ATP, and phosphate in and out of mitochondria. Yet another protein translocator in the inner mitochondrial membrane, the **OXA complex**, mediates the insertion of those inner membrane proteins that are synthesized within mitochondria. It also helps to insert some imported inner membrane proteins that are initially transported into the matrix space by the other complexes.

Mitochondrial Precursor Proteins Are Imported as Unfolded Polypeptide Chains

We have learned almost everything we know about the molecular mechanism of protein import into mitochondria from analyses of cell-free, reconstituted transport systems, in which purified mitochondria in a test tube import radiolabeled mitochondrial precursor proteins. By changing the conditions in the test tube, it is possible to establish the biochemical requirements for the import.

Mitochondrial precursor proteins do not fold into their native structures after they are synthesized; instead, they remain unfolded in the cytosol through interactions with other proteins. Some of these interacting proteins are general *chaperone proteins* of the *Hsp70 family* (discussed in Chapter 6), whereas others are dedicated to mitochondrial precursor proteins and bind directly to their signal sequences. All the interacting proteins help to prevent the precursor proteins from aggregating or folding up spontaneously before they engage with the TOM complex in the outer mitochondrial membrane. As a first step in the import process, the import receptors of the TOM complex bind the signal sequence of the mitochondrial precursor protein. The interacting proteins are then stripped off, and the unfolded polypeptide chain is fed—signal sequence first—into the translocation channel.

In principle, a protein could reach the mitochondrial matrix space by either crossing the two membranes all at once or crossing one at a time. One can distinguish between these possibilities by cooling a cell-free mitochondrial import system to arrest the proteins at an intermediate step in the translocation process. The result is that the arrested proteins no longer contain their N-terminal signal sequence, indicating that the N-terminus must be in the matrix space where the signal peptidase is located, but the bulk of the protein can still be attacked from outside the mitochondria by externally added proteolytic enzymes (**Figure 12–24**). Clearly, the precursor proteins can pass through both mitochondrial membranes at once to enter the matrix space (**Figure 12–25**). It is

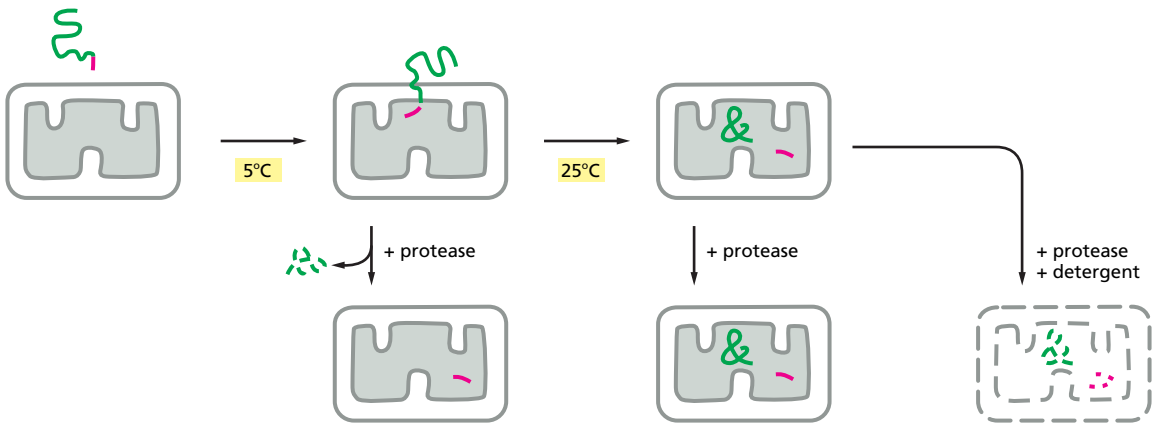


Figure 12–24 Proteins transiently span the inner and outer mitochondrial membranes during their translocation into the matrix space. When isolated mitochondria are incubated with a precursor protein at 5°C, the precursor is only partly translocated across mitochondrial membranes. The N-terminal signal sequence *(red)* is cleaved off in the matrix space; but most of the polypeptide chain remains outside the mitochondrion, where it is accessible to externally added proteolytic enzymes. Upon warming to 25°C, the translocation is completed. Once inside the mitochondrion, the polypeptide chain is protected from added proteolytic enzymes. As a control, when detergents are added to disrupt the mitochondrial membranes, the imported proteins can be readily digested by the same protease treatment.

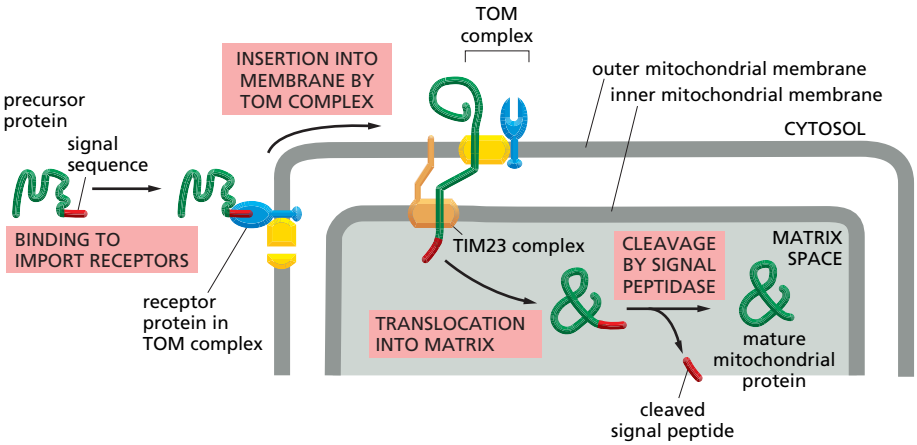


Figure 12–25 Protein import by mitochondria. <ACGG> The N-terminal signal sequence of the mitochondrial precursor protein is recognized by receptors of the TOM complex. The protein is then translocated through the TIM23 complex so that it transiently spans both mitochondrial membranes. The signal sequence is cleaved off by a signal peptidase in the matrix space to form the mature protein. The free signal sequence is then rapidly degraded (not shown).

thought that the TOM complex first transports the signal sequence across the outer membrane to the intermembrane space, where it binds to a TIM complex, opening the channel in the complex. The polypeptide chain then either enters the matrix space or inserts into the inner membrane.

Although the TOM and TIM complexes usually work together to transport precursor proteins across both membranes at the same time, they can work independently. In isolated outer membranes, for example, the TOM complex can translocate the signal sequence of precursor proteins across the membrane. Similarly, if the outer membrane is experimentally disrupted in isolated mitochondria, the exposed TIM23 complex can efficiently import precursor proteins into the matrix space.

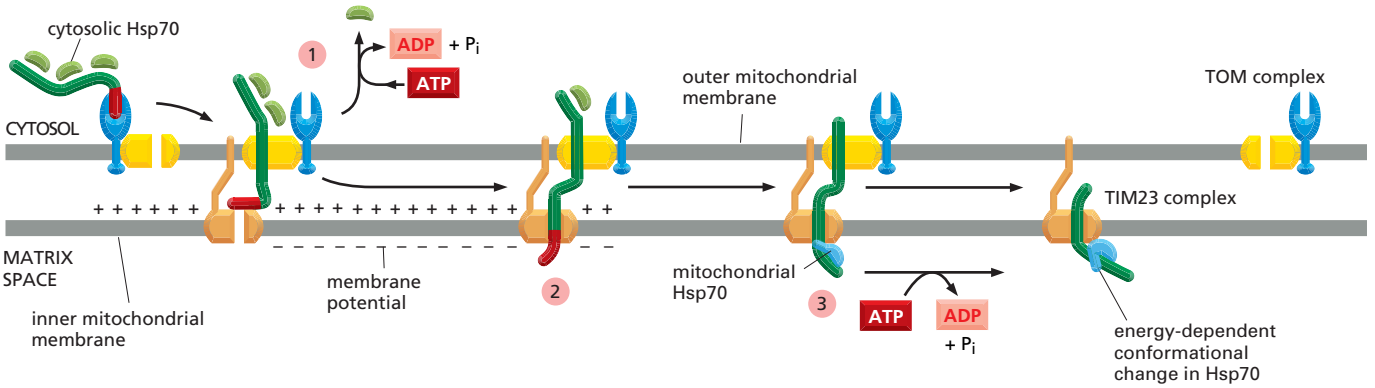
ATP Hydrolysis and a Membrane Potential Drive Protein Import Into the Matrix Space

Directional transport requires energy, which in most biological systems is supplied by ATP hydrolysis. ATP hydrolysis fuels mitochondrial protein import at two discrete sites, one outside the mitochondria and one in the matrix space (**Figure 12–26**). In addition, protein import requires another energy source, which is the membrane potential across the inner mitochondrial membrane.

The first requirement for energy occurs at the initial stage of the translocation process, when the unfolded precursor protein, associated with chaperone proteins, interacts with the import receptors of the TOM complex. As discussed in Chapter 6, the binding and release of newly synthesized polypeptides from the Hsp70 family of chaperone proteins requires ATP hydrolysis. The requirement for Hsp70 and ATP in the cytosol can be bypassed if the precursor protein is artificially unfolded prior to adding it to purified mitochondria.

Once the signal sequence has passed through the TOM complex and is bound to a TIM complex, further translocation through the TIM translocation channel requires the membrane potential, which is the electrical component of the electrochemical H^+ gradient across the inner membrane (see Figure 11–4). Pumping of H^+ from the matrix space to the intermembrane space, driven by

Figure 12–26 The role of energy in protein import into the mitochondrial matrix space. (1) Bound *cytosolic Hsp70* is released from the protein in a step that depends on ATP hydrolysis. After initial insertion of the signal sequence and of adjacent portions of the polypeptide chain into the TOM complex, the signal sequence interacts with a TIM complex. (2) The signal sequence is then translocated into the matrix space in a process that requires a membrane potential across the inner membrane. (3) *Mitochondrial Hsp70*, which is part of an import ATPase complex, binds to regions of the polypeptide chain as they become exposed in the matrix space, pulling the protein through the translocation channel.



electron transport processes in the inner membrane (discussed in Chapter 14), maintains the electrochemical gradient. The energy in the electrochemical H^+ gradient across the inner membrane not only helps drive most of the cell's ATP synthesis, but it also drives the translocation of the positively charged signal sequences through the TIM complexes by electrophoresis.

Mitochondrial Hsp70 also plays a crucial part in the import process. Mitochondria containing mutant forms of the protein fail to import precursor proteins. The Hsp70 is part of a multisubunit protein assembly that is bound to the matrix side of the TIM23 complex and acts as a motor to pull the precursor protein into the matrix space. Like its cytosolic cousin, mitochondrial Hsp70 has a high affinity for unfolded polypeptide chains, and it binds tightly to an imported protein as soon as the protein emerges from the TIM translocator in the matrix space. The Hsp70 then releases the protein in an ATP-dependent step. This energy-driven cycle of binding and subsequent release is thought to provide the final driving force needed to complete protein import after a protein has initially inserted into the TIM23 complex (see Figure 12–26).

After the initial interaction with mitochondrial Hsp70, many imported matrix proteins are passed on to another chaperone protein, *mitochondrial Hsp60*. As discussed in Chapter 6, Hsp60 helps the unfolded polypeptide chain to fold by binding and releasing it through cycles of ATP hydrolysis.

Bacteria and Mitochondria Use Similar Mechanisms to Insert Porins into their Outer Membrane

The outer mitochondrial membrane, like the outer membrane of Gram-negative bacteria (see Figure 11–18), contains abundant pore-forming proteins called porins and is thus freely permeable to inorganic ions and metabolites (but not to most proteins). Porins are β-barrel proteins (see Figure 10–26) and are first imported through the TOM complex (**Figure 12–27**). In contrast to other outer membrane proteins, which are anchored in the membrane through α-helical regions, the TOM complex cannot integrate porins into the lipid bilayer. Instead, porins are first transported into the intermembrane space, where they transiently bind specialized chaperone proteins, which keep the porins from aggregating. They then bind to the SAM complex in the outer membrane, which both inserts them into the outer membrane and helps them fold properly.

One of the central subunits of the SAM complex is homologous to a bacterial outer membrane protein that helps insert β-barrel proteins into the bacterial outer membrane from the periplasmic space (the topological equivalent of the intermembrane space in mitochondria). This conserved pathway for inserting β-barrel proteins is further evidence for the endosymbiotic origin of mitochondria.

Transport Into the Inner Mitochondrial Membrane and Intermembrane Space Occurs Via Several Routes

The same mechanism that transports proteins into the matrix space, using the TOM and TIM23 translocators (see Figure 12–25), also mediates the initial translocation of many proteins that are destined for the inner mitochondrial

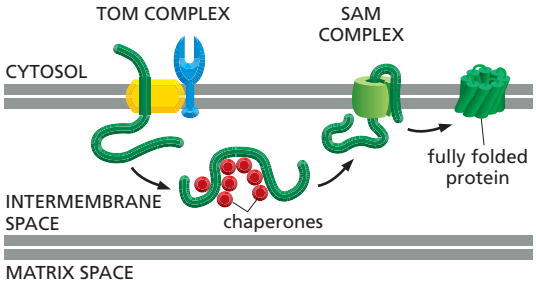


Figure 12–27 Integration of porins into the outer mitochondrial membrane. After translocation through the TOM complex, β-barrel proteins bind to chaperones in the intermembrane space. The SAM complex then inserts the unfolded polypeptide chain into the outer membrane.

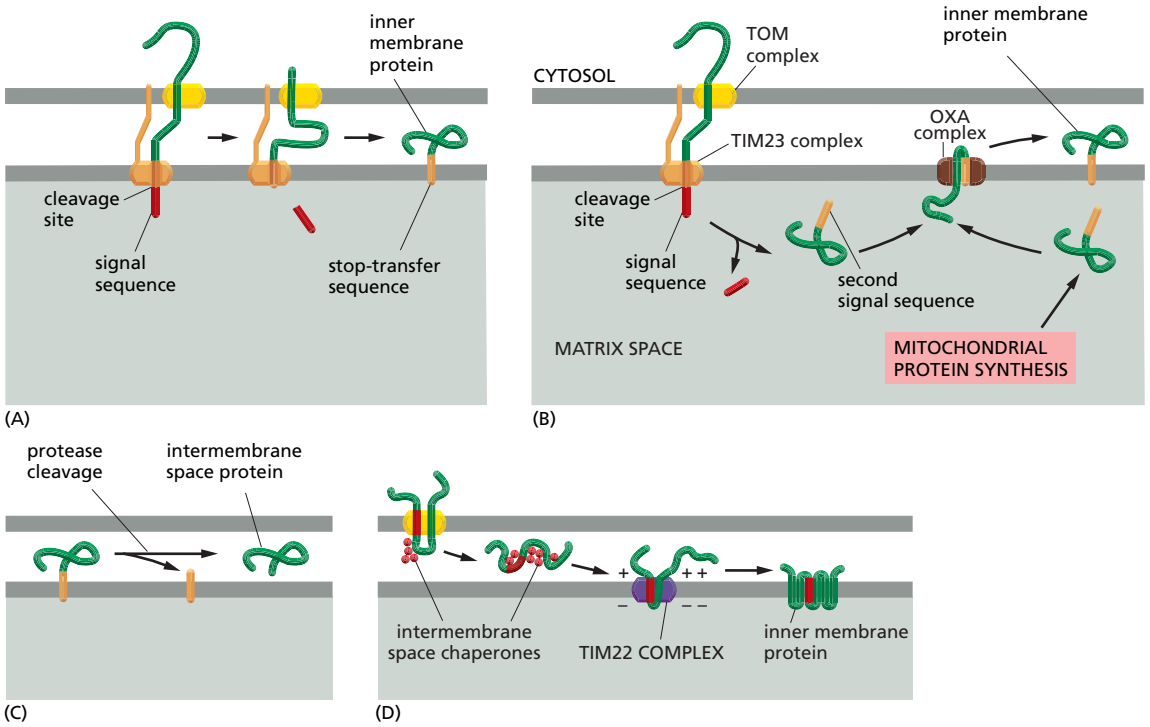


Figure 12–28 Protein import from the cytosol into the inner mitochondrial membrane and intermembrane space. (A) The N-terminal signal sequence (*red*) initiates import into the matrix space (see Figure 12–25). A hydrophobic sequence (*orange*) that follows the matrix-targeting signal binds to the TIM23 translocator in the inner membrane and stops translocation. The remainder of the protein is then pulled into the intermembrane space through the TOM translocator in the outer membrane, and the hydrophobic sequence is released into the inner membrane. (B) A second route for protein integration into the inner membrane first delivers the protein completely into the matrix space. Cleavage of the signal sequence *(red)* used for the initial translocation unmasks an adjacent hydrophobic signal sequence *(orange)* at the new N-terminus. This signal then directs the protein into the inner membrane, using the same OXA-dependent pathway that inserts proteins that are encoded by the mitochondrial genome and translated in the matrix space. (C) Some soluble proteins of the intermembrane space also use the pathways shown in (A) and (B) before they are released into the intermembrane space by a second signal peptidase, which has its active site in the intermembrane space and removes the hydrophobic signal sequence. (D) Metabolite transporters contain internal signal sequences and snake through the TOM complex as a loop. They then bind to the chaperones in the intermembrane space, which guide the proteins to the TIM22 complex. The TIM22 complex is specialized for the insertion of multipass inner membrane proteins.

membrane or the intermembrane space. In the most common translocation route taken, only the N-terminal signal sequence of the transported protein actually enters the matrix space (**Figure 12–28**A). A hydrophobic amino acid sequence, strategically placed after the N-terminal signal sequence, acts as a *stop-transfer sequence,* preventing further translocation across the inner membrane. The TOM complex pulls the remainder of the protein through the outer membrane into the intermembrane space; the signal sequence is cleaved off in the matrix; and the hydrophobic sequence, released from TIM23, remains anchored in the inner membrane.

In another transport route to the inner membrane or intermembrane space, the TIM23 complex initially translocates the entire protein into the matrix space (Figure 12–28B). A matrix signal peptidase then removes the N-terminal signal sequence, exposing a hydrophobic sequence at the new N-terminus. This signal sequence guides the protein to the OXA complex, which inserts the protein into the inner membrane (see Figure 12–23). As mentioned earlier, the OXA complex is primarily used to insert proteins that are encoded and translated in the mitochondrion into the inner membrane, and only a few imported proteins use this pathway. Translocators that are closely related to the OXA complex are found in the plasma membrane of bacteria and in the thylakoid membrane of chloroplasts, where they are thought to help insert membrane proteins by a similar mechanism.

Many proteins that use these pathways to the inner membrane remain anchored there through their hydrophobic signal sequence (see Figure 12–28A and B). Others, however, are released into the intermembrane space by a protease that removes the membrane anchor (Figure 12–28C). Many of these cleaved proteins remain attached to the outer surface of the inner membrane as peripheral subunits of protein complexes that also contain transmembrane proteins.

Mitochondria are the principal sites of ATP synthesis in the cell, but they also contain many metabolic enzymes, such as those of the citric acid cycle. Thus, in addition to proteins, mitochondria must also transport small metabolites across their membranes. While the outer membrane contains porins, which make the membrane freely permeable to such small molecules, the inner membrane does not. Instead, a family of metabolite-specific transporters transfers a vast number of small molecules across the inner membrane. In yeast cells, these transporters comprise a family of 35 different proteins, the most abundant of which transport ATP, ADP, and phosphate. These are multipass transmembrane proteins, which do not have cleavable signal sequences at their N-termini but instead contain internal signal sequences. They cross the TOM complex in the outer membrane, and intermembrane space chaperones guide them to the TIM22 complex, which inserts them into the inner membrane by a process that requires the membrane potential, but not mitochondrial Hsp70 or ATP (Figure 12–28D). An energetically favorable partitioning of the hydrophobic transmembrane regions into the inner membrane is also likely to help drive this process.

Two Signal Sequences Direct Proteins to the Thylakoid Membrane in Chloroplasts

Protein transport into **chloroplasts** resembles transport into mitochondria. Both processes occur post-translationally, use separate translocation complexes in each membrane, require energy, and use amphiphilic N-terminal signal sequences that are removed after use. With the exception of some of the chaperone molecules, however, the protein components that form the translocation complexes differ. Moreover, whereas mitochondria harness the electrochemical H^+ gradient across their inner membrane to drive transport, chloroplasts, which have an electrochemical H^+ gradient across their thylakoid membrane but not their inner membrane, use GTP and ATP hydrolysis to power import across their double membrane. The functional similarities may thus result from convergent evolution, reflecting the common requirements for translocation across a double membrane.

Although the signal sequences for import into chloroplasts superficially resemble those for import into mitochondria, the same plant cells have both mitochondria and chloroplasts, so proteins must partition appropriately between them. In plants, for example, a bacterial enzyme can be directed specifically to mitochondria if it is experimentally joined to an N-terminal signal sequence of a mitochondrial protein; the same enzyme joined to an N-terminal signal sequence of a chloroplast protein ends up in chloroplasts. Thus, the import receptors on each organelle distinguish between the different signal sequences.

Chloroplasts have an extra membrane-enclosed compartment, the **thylakoid**. Many chloroplast proteins, including the protein subunits of the photosynthetic system and of the ATP synthase (discussed in Chapter 14) are located in the thylakoid membrane. Like the precursors of some mitochondrial proteins, the transport of these precursor proteins from the cytosol to their final destination occurs in two steps. First, they pass across the double membrane at special contact sites into the matrix space (called the **stroma** in chloroplasts) and then they translocate either into the thylakoid membrane or into the thylakoid space (**Figure 12–29**A). The precursors of these proteins have a hydrophobic thylakoid signal sequence following the N-terminal chloroplast signal sequence. After the N-terminal signal sequence has been used to import the protein into the stroma,

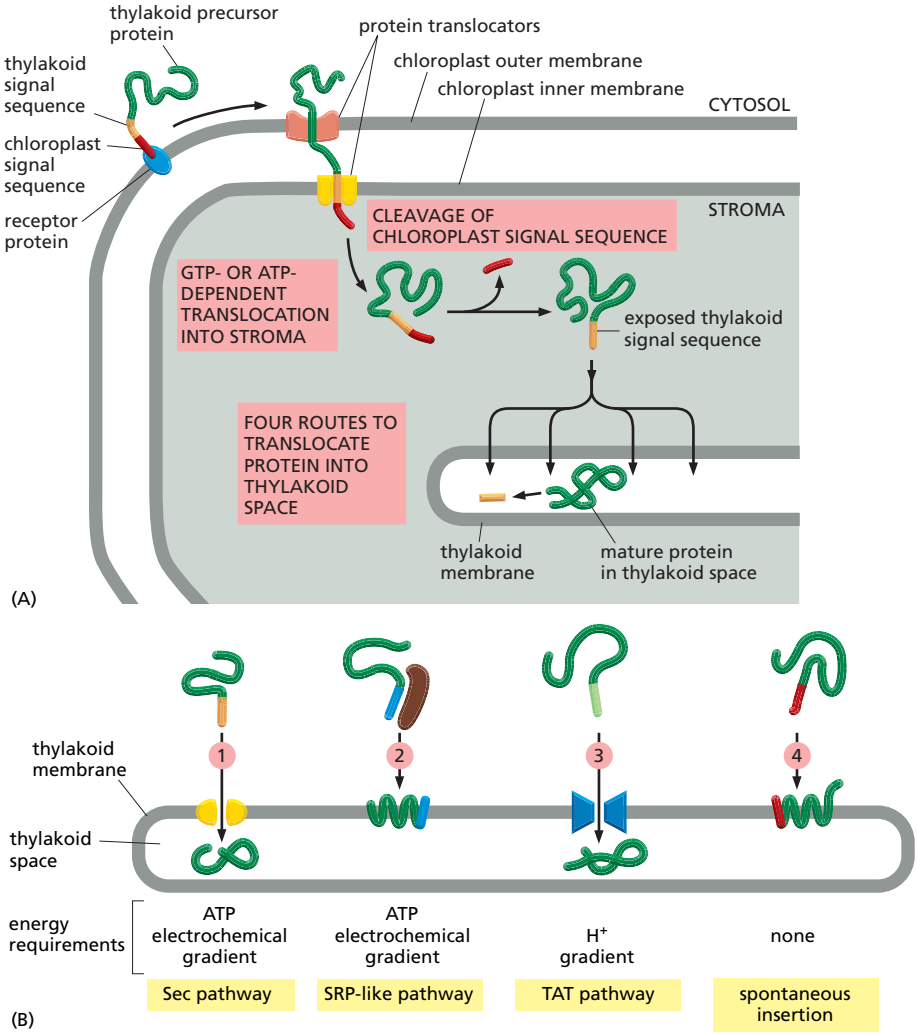


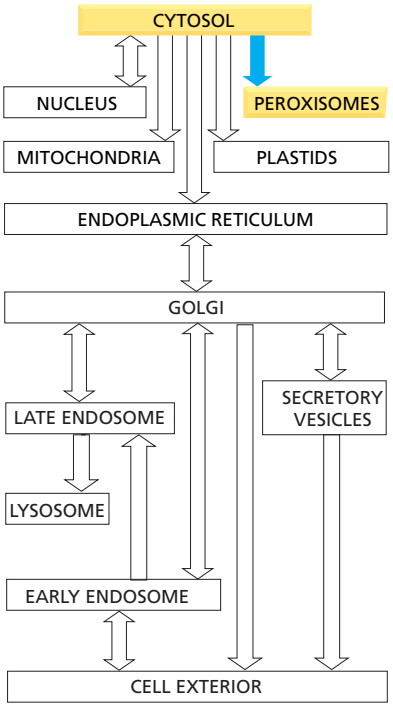
Figure 12–29 Translocation of chloroplast precursor proteins into the thylakoid space. (A) The precursor protein contains an N-terminal chloroplast signal sequence *(red)*, followed immediately by a thylakoid signal sequence *(orange)*. The chloroplast signal sequence initiates translocation into the stroma through a membrane contact site by a mechanism similar to that used for the translocation of mitochondrial precursor proteins into the matrix space. The signal sequence is then cleaved off, unmasking the thylakoid signal sequence, which initiates translocation across the thylakoid membrane. (B) Translocation into the thylakoid space or thylakoid membrane can occur by any one of at least four routes: (1) a *Sec pathway*, so called because it uses components that are homologs of Sec proteins, which mediate protein translocation across the bacterial plasma membrane (discussed later), (2) an *SRP-like pathway*, so called because it uses a chloroplast homolog of the signal recognition particle, or SRP (discussed later); (3) a *TAT* (twin arginine translocation) *pathway*, so called because two arginines are critical in the signal sequences that direct proteins into this pathway, which depends on the H^+ gradient across the thylakoid membrane; and (4) a *spontaneous insertion pathway* that seems not to require any protein translocator.

a stromal signal peptidase removes it, unmasking the thylakoid signal sequence that initiates transport across the thylakoid membrane. There are at least four routes by which proteins cross or become integrated into the thylakoid membrane, distinguished by their need for different stromal chaperones and energy sources (Figure 12–29B).

Summary

Although mitochondria and chloroplasts have their own genetic systems, they produce only a small proportion of their own proteins. Instead, the two organelles import most of their proteins from the cytosol, using similar mechanisms. In both cases, proteins are transported in an unfolded state across both outer and inner membranes simultaneously into the matrix space or stroma. Both ATP hydrolysis and a membrane potential across the inner membrane drive translocation into mitochondria, whereas GTP and ATP hydrolysis drive translocation into chloroplasts. Chaperone proteins of the cytosolic Hsp70 family maintain the precursor proteins in an unfolded state, and a second set of Hsp70 proteins in the matrix space or stroma pull the polypeptide chain into the organelle. Only proteins that contain a specific signal sequence are translocated. The signal sequence can either be located at the N-terminus and cleaved off after import or be internal and retained. Transport into the inner membrane sometimes uses a second, hydrophobic signal sequence that is unmasked when the first signal sequence is removed. In chloroplasts, import from the stroma into the thylakoid can occur by several routes, distinguished by the chaperones and energy source used.

PEROXISOMES



Peroxisomes differ from mitochondria and chloroplasts in many ways. Most notably, they are surrounded by only a single membrane, and they do not contain DNA or ribosomes. Thus, lacking a genome, all of their proteins are encoded in the nucleus. Peroxisomes acquire most of these proteins by selective import from the cytosol, although some of them enter the peroxisome membrane via the ER.

Because we do not discuss peroxisomes elsewhere, we shall digress to consider some of the functions of this diverse family of organelles, before discussing their biosynthesis. All eucaryotic cells have peroxisomes. They contain oxidative enzymes, such as *catalase* and *urate oxidase*, at such high concentrations that, in some cells, the peroxisomes stand out in electron micrographs because of the presence of a crystalloid core (**Figure 12–30**).

Like mitochondria, peroxisomes are major sites of oxygen utilization. One hypothesis is that peroxisomes are a vestige of an ancient organelle that performed all the oxygen metabolism in the primitive ancestors of eucaryotic cells. When the oxygen produced by photosynthetic bacteria first accumulated in the atmosphere, it would have been highly toxic to most cells. Peroxisomes might have lowered the intracellular concentration of oxygen, while also exploiting its chemical reactivity to perform useful oxidation reactions. According to this view, the later development of mitochondria rendered peroxisomes largely obsolete because many of the same biochemical reactions—which had formerly been carried out in peroxisomes without producing energy—were now coupled to ATP formation by means of oxidative phosphorylation. The oxidation reactions performed by peroxisomes in present-day cells would therefore be those whose functions were not taken over by mitochondria.

Peroxisomes Use Molecular Oxygen and Hydrogen Peroxide to Perform Oxidation Reactions

Peroxisomes are so named because they usually contain one or more enzymes that use molecular oxygen to remove hydrogen atoms from specific organic substrates (designated here as R) in an oxidation reaction that produces *hydrogen peroxide (H_2O_2):*

$$RH_2 + O_2 \rightarrow R + H_2O_2$$

Catalase uses the H_2O_2 generated by other enzymes in the organelle to oxidize a variety of other substrates—including phenols, formic acid, formaldehyde, and alcohol—by the "peroxidation" reaction: $H_2O_2 + R'H_2 \rightarrow R' + 2H_2O$. This type of oxidation reaction is particularly important in liver and kidney cells, where the peroxisomes detoxify various toxic molecules that enter the bloodstream. About 25% of the ethanol we drink is oxidized to acetaldehyde in this way. In addition, when excess H_2O_2 accumulates in the cell, catalase converts it to H_2O through the reaction

$$2H_2O_2 \rightarrow 2H_2O + O_2$$

A major function of the oxidation reactions performed in peroxisomes is the breakdown of fatty acid molecules. The process called *β oxidation* shortens the alkyl chains of fatty acids sequentially in blocks of two carbon atoms at a time, thereby converting the fatty acids to acetyl CoA. The peroxisomes then export the acetyl CoA to the cytosol for reuse in biosynthetic reactions. In mammalian cells, β oxidation occurs in both mitochondria and peroxisomes; in yeast and plant cells, however, this essential reaction occurs exclusively in peroxisomes.

An essential biosynthetic function of animal peroxisomes is to catalyze the first reactions in the formation of *plasmalogens*, which are the most abundant class of phospholipids in myelin (**Figure 12–31**). Plasmalogen deficiencies cause profound abnormalities in the myelination of nerve cell axons, which is why many peroxisomal disorders lead to neurological disease.

Peroxisomes are unusually diverse organelles, and even in the various cell types of a single organism they may contain different sets of enzymes. They also adapt remarkably to changing conditions. Yeasts grown on sugar, for example,

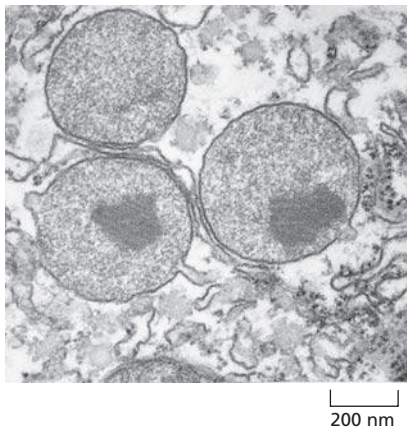


Figure 12–30 An electron micrograph of three peroxisomes in a rat liver cell. The paracrystalline, electron-dense inclusions are composed of the enzyme urate oxidase. (Courtesy of Daniel S. Friend.)

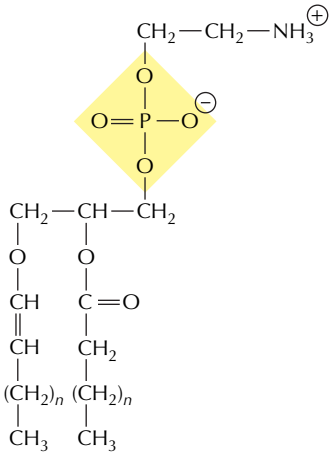


Figure 12–31 The structure of a plasmalogen. Plasmalogens are very abundant in the myelin sheaths that insulate the axons of nerve cells. They make up some 80–90% of the myelin membrane phospholipids. In addition to an ethanolamine head group and a long-chain fatty acid attached to the same glycerol phosphate backbone used for phospholipids, plasmalogens contain an unusual fatty alcohol that is attached through an ether linkage *(bottom left)*.

have small peroxisomes. But when some yeasts are grown on methanol they develop large peroxisomes that oxidize methanol; and when grown on fatty acids they develop large peroxisomes that break down fatty acids to acetyl CoA by β oxidation.

Peroxisomes are also important in plants. Two types of plant peroxisomes have been studied extensively. One is present in leaves, where it participates in *photorespiration* (discussed in Chapter 14) (**Figure 12–32**A). The other type of peroxisome is present in germinating seeds, where it converts the fatty acids stored in seed lipids into the sugars needed for the growth of the young plant. Because this conversion of fats to sugars is accomplished by a series of reactions known as the *glyoxylate cycle,* these peroxisomes are also called *glyoxysomes* (Figure 12–32B). In the glyoxylate cycle, two molecules of acetyl CoA produced by fatty acid breakdown in the peroxisome are used to make succinic acid, which then leaves the peroxisome and is converted into glucose in the cytosol. The glyoxylate cycle does not occur in animal cells, and animals are therefore unable to convert the fatty acids in fats into carbohydrates.

A Short Signal Sequence Directs the Import of Proteins into Peroxisomes

A specific sequence of three amino acids (Ser–Lys–Leu) located at the C-terminus of many peroxisomal proteins functions as an import signal (see Table 12–3, p. 702). Other peroxisomal proteins contain a signal sequence near the N-terminus. If either sequence is attached to a cytosolic protein, the protein is imported into peroxisomes. The import process is still poorly understood, although it is known to involve both soluble receptor proteins in the cytosol, which recognize the targeting signals, and docking proteins on the cytosolic surface of the peroxisome. At least 23 distinct proteins, called **peroxins**, participate in the import process, which is driven by ATP hydrolysis. A complex of at least six different peroxins forms a membrane translocator. Because even oligomeric proteins do not have to unfold to be imported into peroxisomes, the mechanism differs from that used by mitochondria and chloroplasts. At least one soluble import receptor, the peroxin Pex5, accompanies its cargo all the way into peroxisomes and, after cargo release, cycles back to the cytosol. These aspects of peroxisomal protein import resemble protein transport into the nucleus.

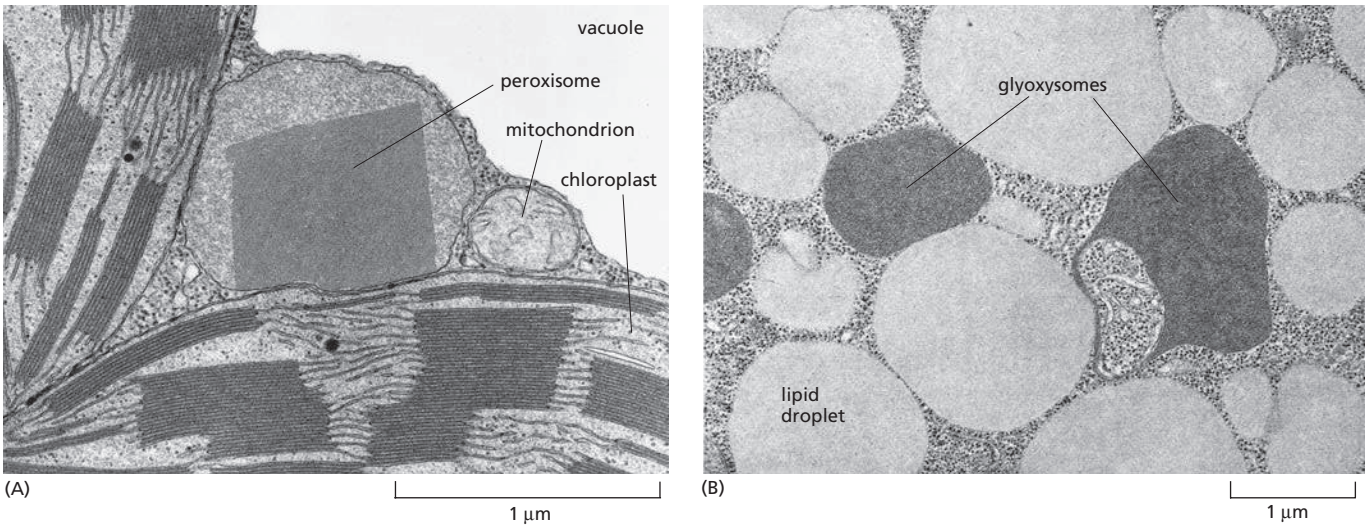


Figure 12–32 Electron micrographs of two types of peroxisomes found in plant cells. (A) A peroxisome with a paracrystalline core in a tobacco leaf mesophyll cell. Its close association with chloroplasts is thought to facilitate the exchange of materials between these organelles during photorespiration. (B) Peroxisomes in a fat-storing cotyledon cell of a tomato seed 4 days after germination. Here the peroxisomes (glyoxysomes) are associated with the lipid droplets that store fat, reflecting their central role in fat mobilization and gluconeogenesis during seed germination. (A, from S.E. Frederick and E.H. Newcomb, *J. Cell Biol.* 43:343–353, 1969. With permission from The Rockefeller Press; B, from W.P. Wergin, P.J. Gruber and E.H. Newcomb, *J. Ultrastruct. Res.* 30:533–557, 1970. With permission from Academic Press.)

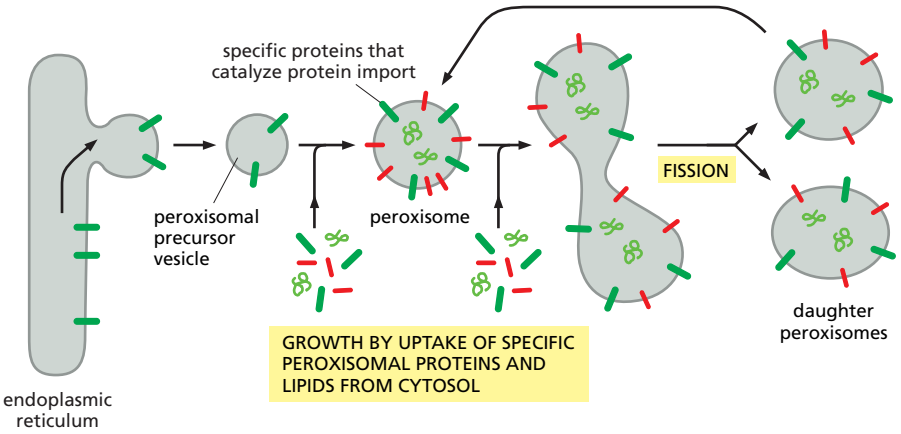


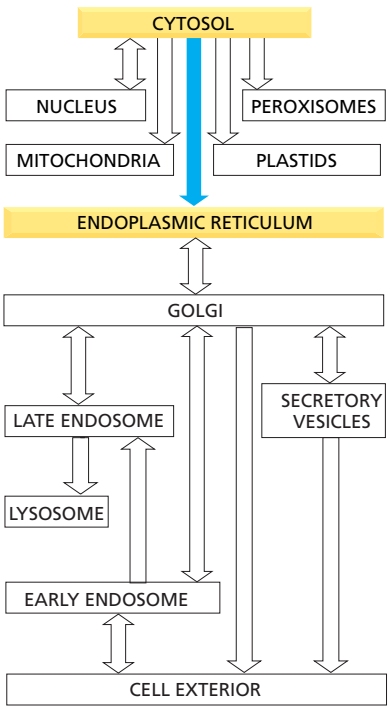
Figure 12–33 A model that explains how peroxisomes proliferate and how new peroxisomes may arise. Peroxisome precursor vesicles are thought to bud from the ER. The machinery that drives the budding reaction and that selects only peroxisomal proteins for packaging into these vesicles—and not ER proteins or proteins destined for other locations in the cell—is not known. Peroxisome precursor vesicles may then fuse with one another or with preexisting peroxisomes. The peroxisome membrane contains import receptor proteins. Cytosolic ribosomes synthesize peroxisomal proteins, including new copies of the import receptor, and then import them into the organelle. Presumably, the lipids required for growth are also imported, although some may derive directly from the ER in the membrane of peroxisome precursor vesicles. (We discuss later the transport of lipids made in the ER through the cytosol to other organelles.)

The importance of this import process and of peroxisomes is demonstrated by the inherited human disease *Zellweger syndrome,* in which a defect in importing proteins into peroxisomes leads to a profound peroxisomal deficiency. These individuals, whose cells contain "empty" peroxisomes, have severe abnormalities in their brain, liver, and kidneys, and they die soon after birth. A mutation in the gene encoding peroxin Pex2, a peroxisomal integral membrane protein involved in protein import, causes one form of the disease. A defective receptor for the N-terminal import signal causes a milder inherited peroxisomal disease.

It has long been debated whether new peroxisomes arise from preexisting ones by organelle growth and fission—and therefore replicate in an autonomous way as mentioned earlier for mitochondria and plastids—or whether they derive as a specialized compartment from the endoplasmic reticulum (ER). Aspects of both views may be true (**Figure 12–33**). Most peroxisomal membrane proteins are made in the cytosol and insert into the membrane of preexisting ones, yet others are first integrated into the ER membrane from where they may bud in specialized peroxisomal precursor vesicles. New precursor vesicles may then fuse with one another and begin importing additional peroxisomal proteins using their own protein import machinery to grow into mature peroxisomes, which can enter into a cycle of growth and fission.

Summary

Peroxisomes are specialized for carrying out oxidation reactions using molecular oxygen. They generate hydrogen peroxide, which they employ for oxidative purposes—and contain catalase to destroy the excess. Like mitochondria and plastids, peroxisomes are self-replicating organelles. Because they do not contain DNA or ribosomes, however, all of their proteins are encoded in the cell nucleus. Some of these proteins are conveyed to peroxisomes via the ER, but most are synthesized in the cytosol. A specific sequence of three amino acids near the C-terminus of many of the cytosolic proteins functions as a peroxisomal import signal. The mechanism of protein import differs from that of mitochondria and chloroplasts, in that even oligomeric proteins are imported from the cytosol without unfolding.



THE ENDOPLASMIC RETICULUM

All eucaryotic cells have an **endoplasmic reticulum (ER)**. Its membrane typically constitutes more than half of the total membrane of an average animal cell (see Table 12–2, p. 697). The ER is organized into a netlike labyrinth of branching tubules and flattened sacs that extends throughout the cytosol (**Figure 12–34**). The tubules and sacs interconnect, and their membrane is continuous with the outer nuclear membrane. Thus, the ER and nuclear membranes form a continuous sheet enclosing a single internal space, called the **ER lumen** or the *ER cisternal space,* which often occupies more than 10% of the total cell volume (see Table 12–1, p. 697).

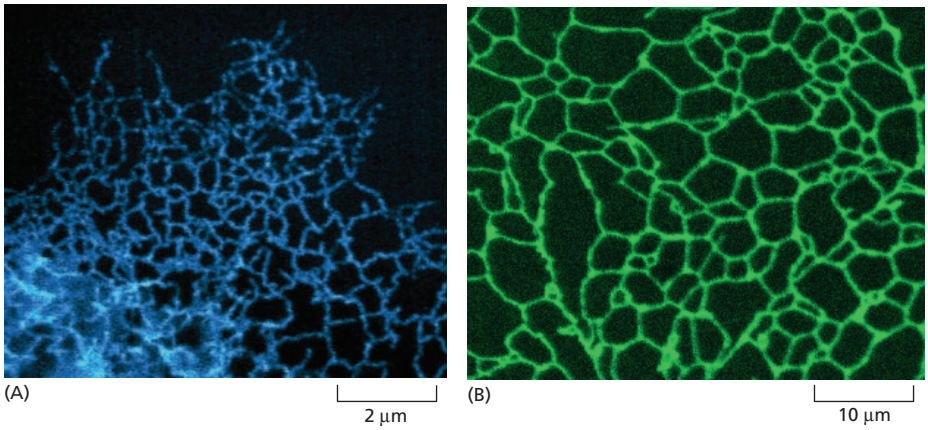


Figure 12–34 Fluorescent micrographs of the endoplasmic reticulum. <TCGT> (A) Part of the ER network in a cultured mammalian cell, stained with an antibody that binds to a protein retained in the ER. The ER extends as a network throughout the entire cytosol, so that all regions of the cytosol are close to some portion of the ER membrane. (B) Part of an ER network in a living plant cell that was genetically engineered to express a fluorescent protein in the ER. (A, courtesy of Hugh Pelham; B, courtesy of Petra Boevink and Chris Hawes.)

The ER has a central role in both lipid and protein biosynthesis, and it also serves as an intracellular Ca^{2+} store that is used in many cell signaling responses (discussed in Chapter 15). The ER membrane is the site of production of all the transmembrane proteins and lipids for most of the cell's organelles, including the ER itself, the Golgi apparatus, lysosomes, endosomes, secretory vesicles, and the plasma membrane. The ER membrane also makes most of the lipids for mitochondrial and peroxisomal membranes. In addition, almost all of the proteins that will be secreted to the cell exterior—plus those destined for the lumen of the ER, Golgi apparatus, or lysosomes—are initially delivered to the ER lumen.

The ER Is Structurally and Functionally Diverse

While the various functions of the ER are essential to every cell, their relative importance varies greatly between individual cell types. To meet different functional demands, distinct regions of the ER become highly specialized. We observe such functional specialization as dramatic changes in ER structure, and different cell types can therefore possess characteristically different types of ER membrane. One of the most remarkable ER specializations is the *rough ER*.

Mammalian cells begin to import most proteins into the ER before complete synthesis of the polypeptide chain—that is, import is a **co-translational** process (**Figure 12–35**A). In contrast, the import of proteins into mitochondria, chloroplasts, nuclei, and peroxisomes is a **post-translational** process (Figure 12–35B). In co-translational transport, the ribosome that is synthesizing the protein is attached directly to the ER membrane, enabling one end of the protein to be translocated into the ER while the rest of the polypeptide chain is being assembled. These membrane-bound ribosomes coat the surface of the ER, creating regions termed **rough endoplasmic reticulum**, or **rough ER** (**Figure 12–36**A).

Regions of ER that lack bound ribosomes are called **smooth endoplasmic reticulum**, or **smooth ER**. Most cells have scanty regions of smooth ER, and the

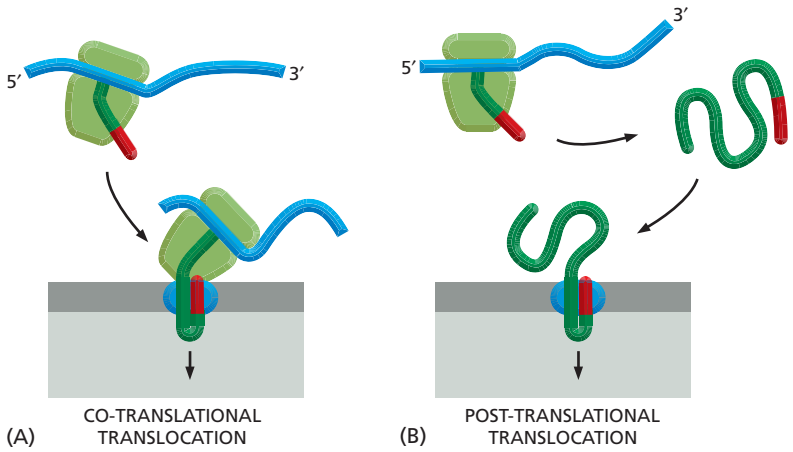


Figure 12–35 Co-translational and post-translational protein translocation. (A) Ribosomes bind to the ER membrane during co-translational translocation. (B) By contrast, ribosomes complete the synthesis of a protein and release it prior to post-translational translocation.

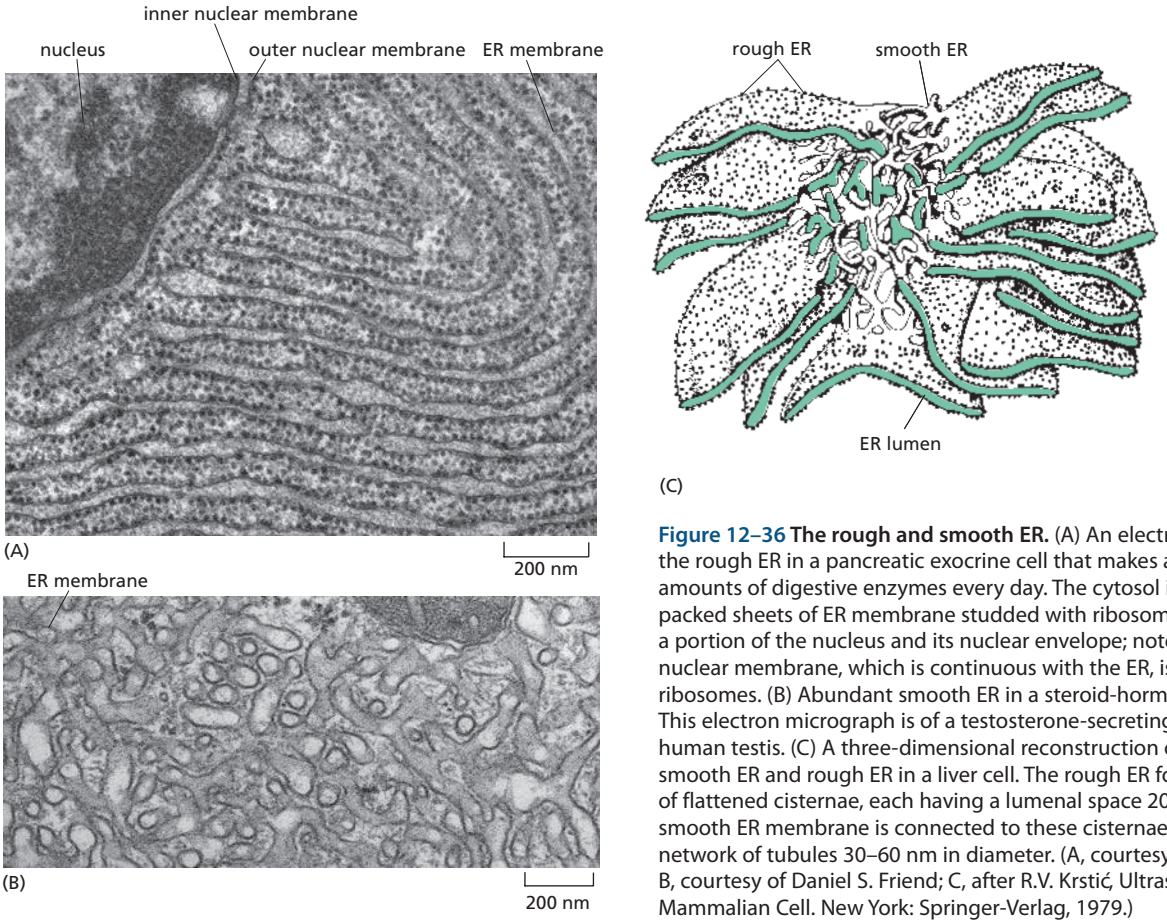


Figure 12–36 The rough and smooth ER. (A) An electron micrograph of the rough ER in a pancreatic exocrine cell that makes and secretes large amounts of digestive enzymes every day. The cytosol is filled with closely packed sheets of ER membrane studded with ribosomes. At the top left is a portion of the nucleus and its nuclear envelope; note that the outer nuclear membrane, which is continuous with the ER, is also studded with ribosomes. (B) Abundant smooth ER in a steroid-hormone-secreting cell. This electron micrograph is of a testosterone-secreting Leydig cell in the human testis. (C) A three-dimensional reconstruction of a region of smooth ER and rough ER in a liver cell. The rough ER forms oriented stacks of flattened cisternae, each having a lumenal space 20–30 nm wide. The smooth ER membrane is connected to these cisternae and forms a fine network of tubules 30–60 nm in diameter. (A, courtesy of Lelio Orci; B, courtesy of Daniel S. Friend; C, after R.V. Krstić, Ultrastructure of the Mammalian Cell. New York: Springer-Verlag, 1979.)

ER is often partly smooth and partly rough. Areas of smooth ER from which transport vesicles carrying newly synthesized proteins and lipids bud off for transport to the Golgi apparatus are called *transitional ER*. In certain specialized cells, the smooth ER is abundant and has additional functions. It is prominent, for example, in cells that specialize in lipid metabolism, such as cells that synthesize steroid hormones from cholesterol; the expanded smooth ER accommodates the enzymes that make cholesterol and modify it to form the hormones (Figure 12–36B).

The main cell type in the liver, the *hepatocyte*, also has abundant smooth ER. It is the principal site of production of *lipoprotein particles*, which carry lipids via the bloodstream to other parts of the body. The enzymes that synthesize the lipid components of the particles are located in the membrane of the smooth ER, which also contains enzymes that catalyze a series of reactions to detoxify both lipid-soluble drugs and various harmful compounds produced by metabolism. The most extensively studied of these *detoxification reactions* are carried out by the *cytochrome P450* family of enzymes, which catalyze a series of reactions in which water-insoluble drugs or metabolites that would otherwise accumulate to toxic levels in cell membranes are rendered sufficiently water-soluble to leave the cell and be excreted in the urine. Because the rough ER alone cannot house enough of these and other necessary enzymes, a major portion of the membrane in a hepatocyte normally consists of smooth ER (Figure 12–36C; see Table 12–2).

Another crucially important function of the ER in most eucaryotic cells is to sequester Ca^{2+} from the cytosol. The release of Ca^{2+} into the cytosol from the ER, and its subsequent reuptake, occurs in many rapid responses to extracellular signals, as discussed in Chapter 15. A Ca^{2+} pump transports Ca^{2+} from the cytosol into the ER lumen. A high concentration of Ca^{2+}-binding proteins in the ER facilitates Ca^{2+} storage. In some cell types, and perhaps in most, specific regions of the ER are specialized for Ca^{2+} storage. Muscle cells have an abundant, modified smooth ER, called the *sarcoplasmic reticulum*. The release and

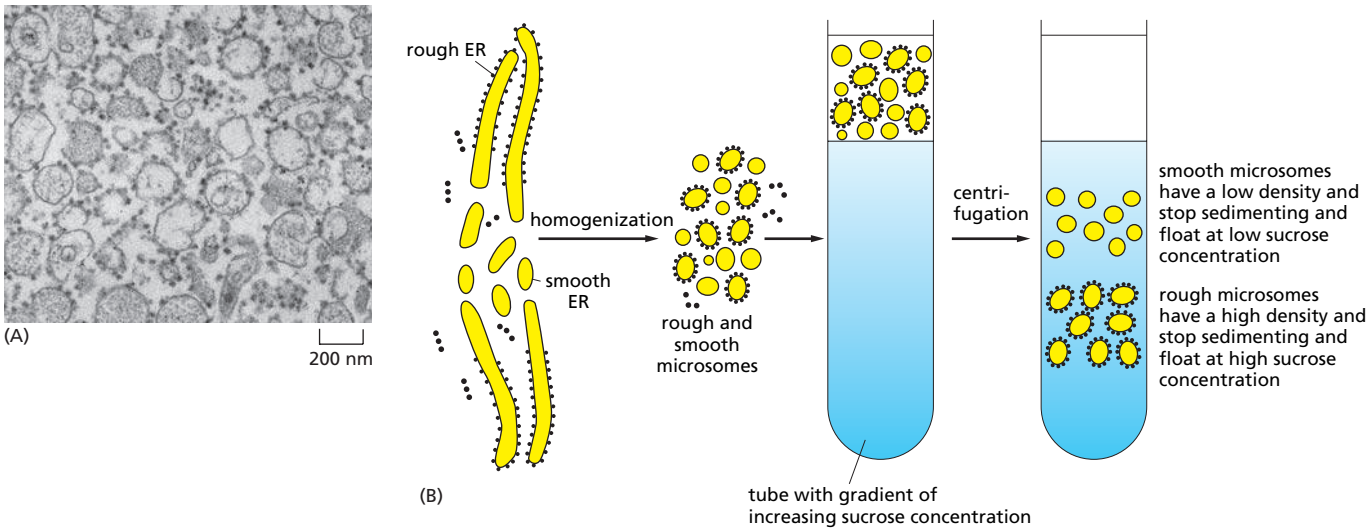


Figure 12–37 The isolation of purified rough and smooth microsomes from the ER. (A) A thin section electron micrograph of the purified rough ER fraction shows an abundance of ribosome-studded vesicles. (B) When sedimented to equilibrium through a gradient of sucrose, the two types of microsomes separate from each other on the basis of their different densities. (A, courtesy of George Palade.)

reuptake of Ca^{2+} by the sarcoplasmic reticulum trigger myofibril contraction and relaxation, respectively, during each round of muscle contraction (discussed in Chapter 16).

To study the functions and biochemistry of the ER, it is necessary to isolate it. This may seem to be a hopeless task because the ER is intricately interleaved with other components of the cytoplasm. Fortunately, when tissues or cells are disrupted by homogenization, the ER breaks into fragments, which reseal to form small (~100–200 nm in diameter) closed vesicles called **microsomes**. Microsomes are relatively easy to purify. To the biochemist, microsomes represent small authentic versions of the ER, still capable of protein translocation, protein glycosylation, Ca^{2+} uptake and release, and lipid synthesis. Microsomes derived from rough ER are studded with ribosomes and are called *rough microsomes*. The ribosomes are always found on the outside surface, so the interior of the microsome is biochemically equivalent to the lumenal space of the ER (**Figure 12–37**A).

Many vesicles similar in size to rough microsomes, but lacking attached ribosomes, are also found in cell homogenates. Such *smooth microsomes* are derived in part from smooth portions of the ER and in part from vesiculated fragments of the plasma membrane, Golgi apparatus, endosomes, and mitochondria (the ratio depending on the tissue). Thus, whereas rough microsomes are clearly derived from rough portions of ER, it is not easy to determine the origins of "smooth microsomes" prepared from disrupted cells. The smooth microsomes prepared from liver or muscle cells are an exception. Because of the unusually large quantities of smooth ER or sarcoplasmic reticulum, respectively, most of the smooth microsomes in homogenates of these tissues are derived from the smooth ER. The ribosomes attached to rough microsomes make them more dense than smooth microsomes. As a result, we can use equilibrium centrifugation to separate the rough and smooth microsomes (Figure 12–37B). Microsomes have been invaluable in elucidating the molecular aspects of ER function, as we discuss next.

Signal Sequences Were First Discovered in Proteins Imported into the Rough ER

The ER captures selected proteins from the cytosol as they are being synthesized. These proteins are of two types: *transmembrane proteins,* which are only partly translocated across the ER membrane and become embedded in it, and *water-soluble proteins,* which are fully translocated across the ER membrane and are released into the ER lumen. Some of the transmembrane proteins function in the ER, but many are destined to reside in the plasma membrane or the membrane of another organelle. The water-soluble proteins are destined either for secretion or for residence in the lumen of an organelle. All of these proteins,

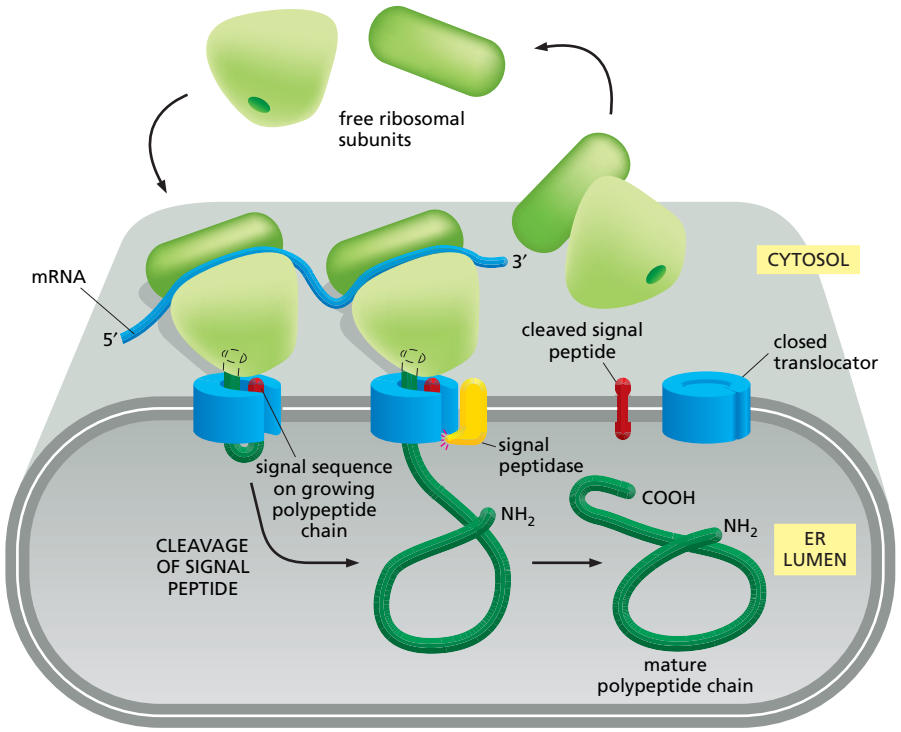


Figure 12–38 The signal hypothesis. A simplified view of protein translocation across the ER membrane, as originally proposed. When the ER signal sequence emerges from the ribosome, it directs the ribosome to a translocator on the ER membrane that forms a pore in the membrane through which the polypeptide is translocated. A signal peptidase is closely associated with the translocator and clips off the signal sequence during translation, and the mature protein is released into the lumen of the ER immediately after synthesis. The translocator is closed until the ribosome has bound, so that the permeability barrier of the ER membrane is maintained at all times.

regardless of their subsequent fate, are directed to the ER membrane by an **ER signal sequence**, which initiates their translocation by a common mechanism.

Signal sequences (and the signal sequence strategy of protein sorting) were first discovered in the early 1970s in secreted proteins that are translocated across the ER membrane as a first step toward their eventual discharge from the cell. In the key experiment, the mRNA encoding a secreted protein was translated by ribosomes *in vitro*. When microsomes were omitted from this cell-free system, the protein synthesized was slightly larger than the normal secreted protein, the extra length being the N-terminal *leader peptide*. In the presence of microsomes derived from the rough ER, however, a protein of the correct size was produced. According to the *signal hypothesis*, the leader is a signal sequence that directs the secreted protein to the ER membrane and is then cleaved off by a *signal peptidase* in the ER membrane before the polypeptide chain has been completed (**Figure 12–38**). Cell-free systems in which proteins are imported into microsomes have provided powerful procedures for identifying, purifying, and studying the various components of the molecular machinery responsible for the ER import process.

A Signal-Recognition Particle (SRP) Directs ER Signal Sequences to a Specific Receptor in the Rough ER Membrane

The ER signal sequence is guided to the ER membrane by at least two components: a **signal-recognition particle** (**SRP**), which cycles between the ER membrane and the cytosol and binds to the signal sequence, and an **SRP receptor** in the ER membrane. The SRP is a complex particle, consisting of six different polypeptide chains bound to a single small RNA molecule (**Figure 12–39**). SRP and its receptor are found in all cells, indicating that this protein-targeting mechanism arose early in evolution and has been conserved.

ER signal sequences vary greatly in amino acid sequence, but each has eight or more nonpolar amino acids at its center (see Table 12–3, p. 702). How can the SRP bind specifically to so many different sequences? The answer has come from the crystal structure of the SRP protein, which shows that the signal-sequence-binding site is a large hydrophobic pocket lined by methionines. Because methionines have unbranched, flexible side chains, the pocket

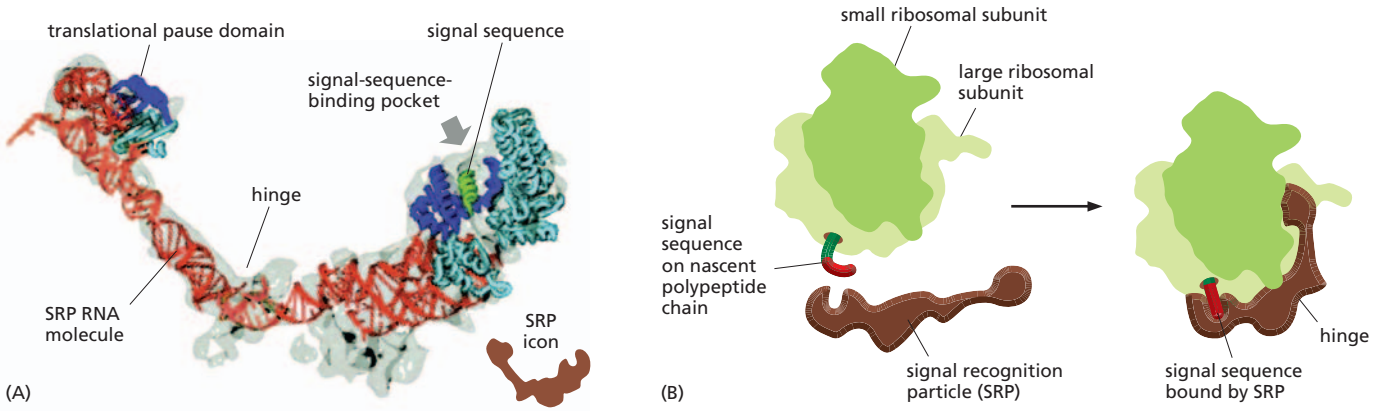


Figure 12–39 The signal-recognition particle (SRP). (A) A mammalian SRP is a rodlike complex containing six protein subunits and one RNA molecule (shown in *red*). The SRP RNA forms the backbone that links the domain of the SRP containing the signal sequence binding pocket to the domain responsible for pausing translation. The three-dimensional outline of the SRP (shown in *gray*) was determined by cryo-electron microscopy. Where known, crystal structures of individual SRP pieces are fitted into the envelope and shown as ribbon diagrams. A bound signal sequence is shown as a *green* helix. (B) SRP bound to the ribsosome visualized by cryo-electron microscopy. SRP binds to the large ribosomal subunit so that its signal sequence binding pocket is positioned near the nascent chain exit site and its translational pause domain is positioned at the interface between the ribosomal subunits, where it interferes with elongation factor binding. (Adapted from M. Halic et al., *Nature* 427:808–814, 2004. With permission from Macmillan Publishers Ltd.)

is sufficiently plastic to accommodate hydrophobic signal sequences of different sequences, sizes, and shapes.

The SRP is a rodlike structure that wraps around the large ribosomal subunit, with one end binding to the ER signal sequence as it emerges as part of the newly made polypeptide chain from the ribosome; the other end blocks the elongation factor binding site at the interface between the large and small ribosomal subunits (Figure 12–39). This block halts protein synthesis as soon as the signal peptide has emerged from the ribosome. The transient pause presumably gives the ribosome enough time to bind to the ER membrane before completion of the polypeptide chain, thereby ensuring that the protein is not released into the cytosol. This safety device may be especially important for secreted and lysosomal hydrolases that could wreak havoc in the cytosol; cells that secrete large amounts of hydrolases, however, take the added precaution of having high concentrations of hydrolase inhibitors in their cytosol. The pause also ensures that large portions of a protein that could fold into a compact structure are not made before reaching the translocator in the ER membrane. Thus, in contrast to the post-translational import of proteins into mitochondria and chloroplasts, chaperone proteins are not required to keep the protein unfolded.

Once formed, the SRP–ribosome complex binds to the SRP receptor, which is an integral membrane protein complex embedded in the rough ER membrane. This interaction brings the SRP–ribosome complex to a protein translocator. The SRP and SRP receptor are then released, and the translocator transfers the growing polypeptide chain across the membrane (**Figure 12–40**).

This co-translational transfer process creates two spatially separate populations of ribosomes in the cytosol. **Membrane-bound ribosomes**, attached to the cytosolic side of the ER membrane, are engaged in the synthesis of proteins that are being concurrently translocated into the ER. **Free ribosomes**, unattached to any membrane, synthesize all other proteins encoded by the nuclear genome. Membrane-bound and free ribosomes are structurally and functionally identical. They differ only in the proteins they are making at any given time.

Since many ribosomes can bind to a single mRNA molecule, a *polyribosome* is usually formed, which becomes attached to the ER membrane, directed there by the signal sequences on multiple growing polypeptide chains (**Figure 12–41**A). The individual ribosomes associated with such an mRNA molecule can return to the cytosol when they finish translation and intermix with the pool of

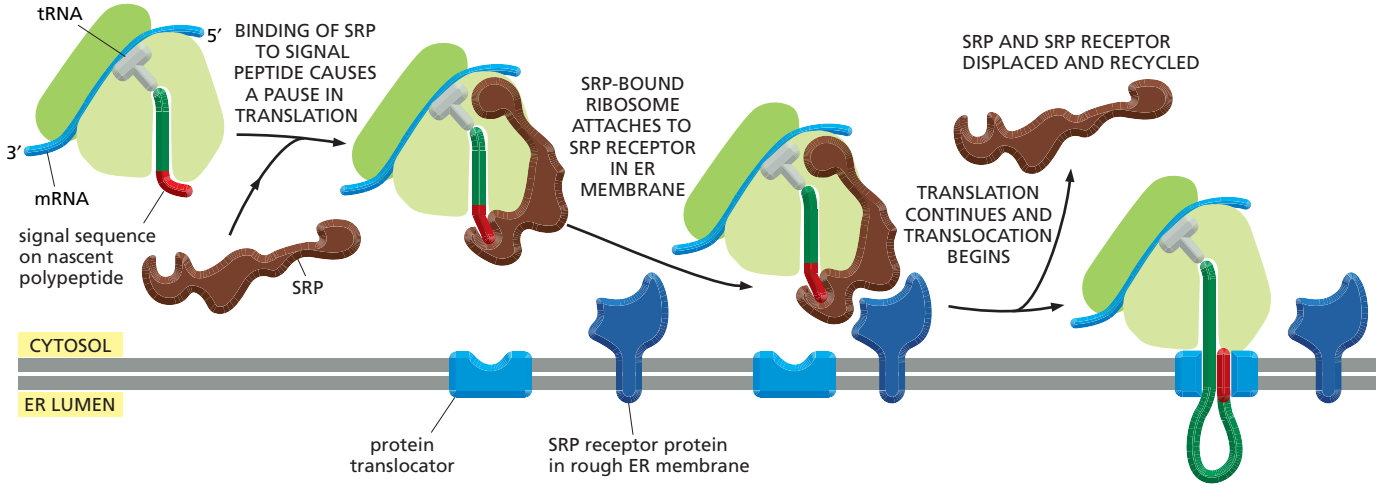
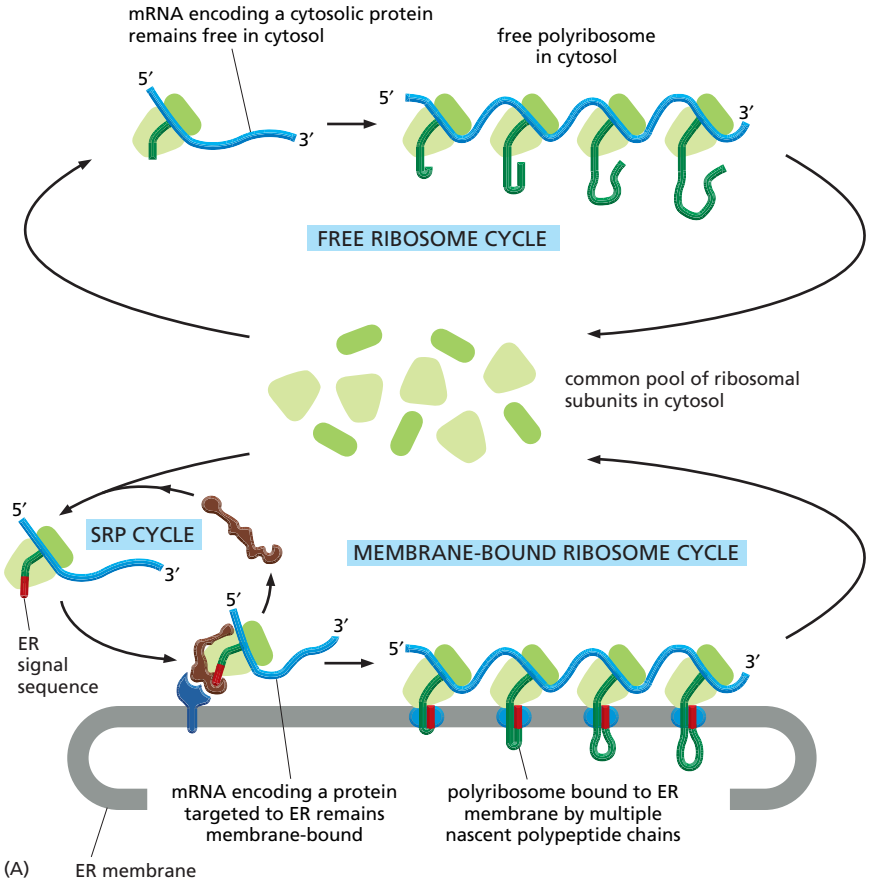


Figure 12–40 How ER signal sequences and SRP direct ribosomes to the ER membrane. <TTCC> The SRP and its receptor are thought to act in concert. The SRP binds to both the exposed ER signal sequence and the ribosome, thereby inducing a pause in translation. The SRP receptor in the ER membrane, which is composed of two different polypeptide chains, binds the SRP–ribosome complex and directs it to the translocator. In a poorly understood reaction, the SRP and SRP receptor are then released, leaving the ribosome bound to the translocator in the ER membrane. The translocator then inserts the polypeptide chain into the membrane and transfers it across the lipid bilayer. Because one of the SRP proteins and both chains of the SRP receptor contain GTP-binding domains, it is thought that conformational changes that occur during cycles of GTP binding and hydrolysis (discussed in Chapter 15) ensure that SRP release occurs only after the ribosome has become properly engaged with the translocator in the ER membrane. The translocator is closed until the ribosome has bound, so that the permeability barrier of the ER membrane is maintained at all times.



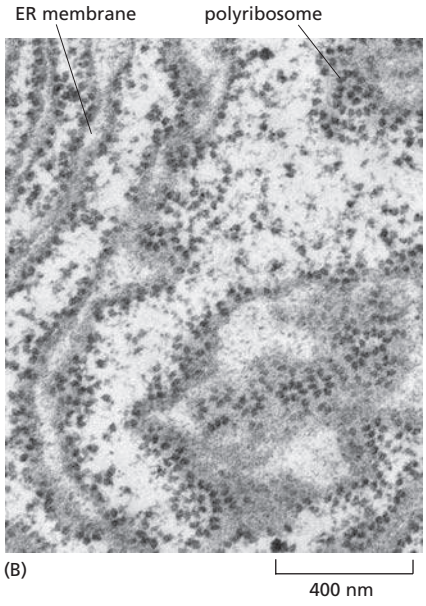


Figure 12–41 Free and membrane-bound ribosomes. (A) A common pool of ribosomes synthesizes the proteins that stay in the cytosol and those that are transported into the ER. The ER signal sequence on a newly formed polypeptide chain binds to SRP, which directs the translating ribosome to the ER membrane. The mRNA molecule remains permanently bound to the ER as part of a polyribosome, while the ribosomes that move along it are recycled; at the end of each round of protein synthesis the ribosomal subunits are released and rejoin the common pool in the cytosol. (B) A thin-section electron micrograph of polyribosomes attached to the ER membrane. The plane of section in some places cuts through the ER roughly parallel to the membrane, giving a face-on view of the rosettelike pattern of the polyribosomes. (B, courtesy of George Palade.)

free ribosomes. The mRNA itself, however, remains attached to the ER membrane by a changing population of ribosomes, each transiently held at the membrane by the translocator (Figure 12–41B).

The Polypeptide Chain Passes Through an Aqueous Pore in the Translocator

It has long been debated whether polypeptide chains are transferred across the ER membrane in direct contact with the lipid bilayer or through a pore in a protein translocator. The debate ended with the identification of the translocator, which was shown to form a water-filled pore in the membrane through which the polypeptide chain passes. The core of the translocator, called the **Sec61 complex**, is built from three subunits that are highly conserved from bacteria to eucaryotic cells. Recently, the structure of the Sec61 complex was determined by x-ray crystallography. The structure suggests that α helices contributed by the largest subsunit surround a central pore through which the polypeptide chain may traverse the membrane (**Figure 12–42**). The pore is gated by a short helix that is thought to keep the translocator closed when it is idle and to move aside when it is engaged in passing a polypeptide chain. According to this view, the pore is a dynamic gated structure that opens only transiently when a polypeptide chain traverses the membrane. In an idle translocator it is important to keep the pore closed, so that the membrane remains impermeable to ions, such as Ca^{2+}, which otherwise would leak out of the ER.

The structure of the Sec61 complex suggests that the pore can also open along a seam on its side. This opening allows a translocating peptide chain lateral access into the hydrophobic core of the membrane, a process that is important both for the release of a cleaved signal peptide into the membrane (see Figure 12–38) and for the integration of membrane proteins into the bilayer, as we discuss later.

In eucaryotic cells, four Sec61 complexes form a large translocator assembly that can be visualized on the ribosomes after detergent solubilization of ER membranes (**Figure 12–43**). It is likely that not all four Sec61 complexes in the eucaryotic translocator participate in protein translocation directly. Some of them may be inactive, providing binding sites for the ribosome and for accessory proteins that help polypeptide chains fold as they enter the ER. According to one view, the bound ribosome forms a tight seal with the translocator, with the space inside the ribosome continuous with the lumen of the ER so that no molecules can escape from the ER. Alternatively, the structure of the Sec61 complex suggests that the pore in the translocator could form a tight-fitting diaphragm around the translocating chain that prevents the escape of other molecules.

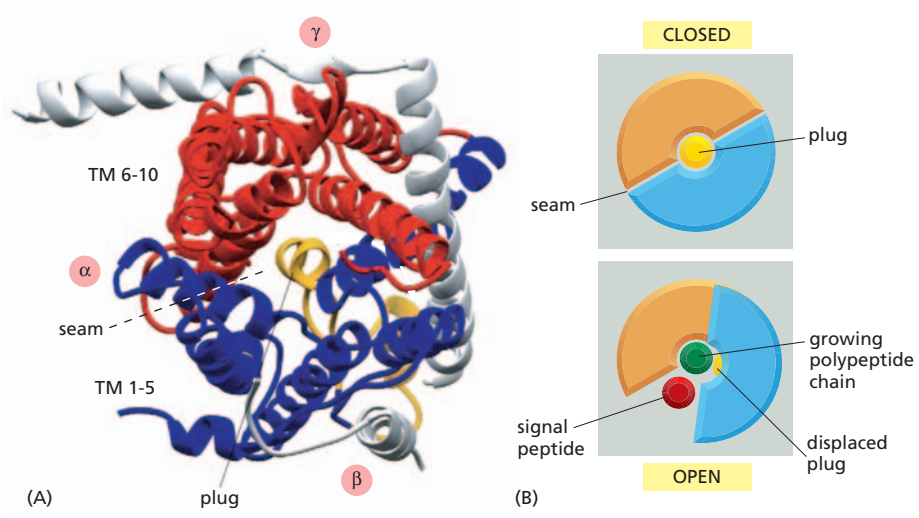


Figure 12–42 Structure of the Sec61 complex. (A) A top view of the structure of the Sec61 complex of the archae *Methanococcus jannaschii* seen from the cytosol side of the membrane. The Sec61α subunit is shown in *blue* and *red;* the two smaller β- and γ-subunits are shown in *gray.* The *yellow* short helix is thought to form a plug that seals the pore when the translocator is closed. To open, the complex rearranges itself to move the plug helix out of the way. It is thought that the pore of the Sec61 complex can also open sideways at a seam. (B) Model of the closed and open states of the translocator are shown, illustrating how a signal sequence could be released into the membrane after opening of the seam. (A, from B. van der Berg et al., *Nature* 427:36–44, 2004. With permission from Macmillan Publishers Ltd.)

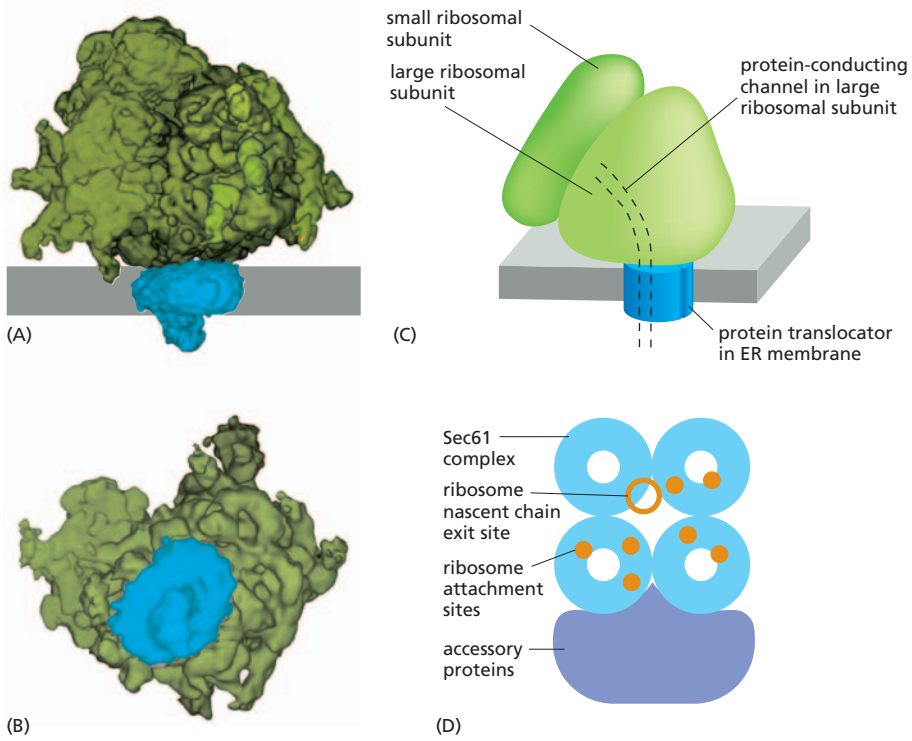


Figure 12–43 A ribosome bound to the eucaryotic protein translocator. (A) A reconstruction of the complex from electron microscopic images viewed from the side. (B) A view of the translocator seen from the ER lumen. The translocator is thought to contain four copies of the Sec61 complex. (C) A schematic drawing of a membrane-bound ribosome attached to the translocator with the tunnel in the large ribosomal subunit, through which the growing polypeptide chain exits from the ribosome (see Figure 6–70). (D) A schematic model of how four Sec61 complexes and various accessory proteins may be arranged in the translocator. The *red dots* indicate ribosome attachment sites that are seen in the electron microscopic reconstructions and the *red circle* shows the position of the end of the ribosome exit tunnel. It is not known which of the four Sec61 complexes are actively participating in the protein translocation event. Domains of the accessory proteins extend across the membrane and form the lumenal bulge seen in the side view in (A). (A, B, and C, adapted from J.F. Ménétret et al., *J. Mol. Biol.* 348:445–457, 2005. With permission from Academic Press.)

Translocation Across the ER Membrane Does Not Always Require Ongoing Polypeptide Chain Elongation

As we have seen, translocation of proteins into mitochondria, chloroplasts, and peroxisomes occurs post-translationally, after the protein has been made and released into the cytosol, whereas translocation across the ER membrane usually occurs during translation (co-translationally). This explains why ribosomes are bound to the ER but usually not to other organelles.

Some completely synthesized proteins, however, are imported into the ER, demonstrating that translocation does not always require ongoing translation. Post-translational protein translocation is especially common across the yeast ER membrane and the bacterial plasma membrane (which is thought to be evolutionarily related to the ER; see Figure 12–4). To function in post-translational translocation, the translocator needs accessory proteins that feed the polypeptide chain into the pore and drive translocation (**Figure 12–44**). In bacteria, a translocation motor protein, the *SecA ATPase*, attaches to the cytosolic side of the translocator, where it undergoes cyclic conformational changes driven by ATP hydrolysis. Each time an ATP is hydrolyzed, a portion of the SecA protein inserts into the pore of the translocator, pushing a short segment of the passenger protein with it. As a result of this ratchet mechanism, the SecA protein progressively pushes the polypeptide chain of the transported protein across the membrane.

Eucaryotic cells use a different set of accessory proteins that associate with the Sec61 complex. These proteins span the ER membrane and use a small domain on the lumenal side of the ER membrane to deposit a Hsp70-like chaperone protein (called *BiP*, for *bi*nding *p*rotein) onto the polypeptide chain as it emerges from the pore into the ER lumen. Cycles of BiP binding and release drive unidirectional translocation, as described earlier for the mitochondrial Hsp70 proteins that pull proteins across mitochondrial membranes.

Proteins that are transported into the ER by a post-translational mechanism are first released into the cytosol, where they bind to chaperone proteins to prevent folding, as discussed earlier for proteins destined for mitochondria and chloroplasts.

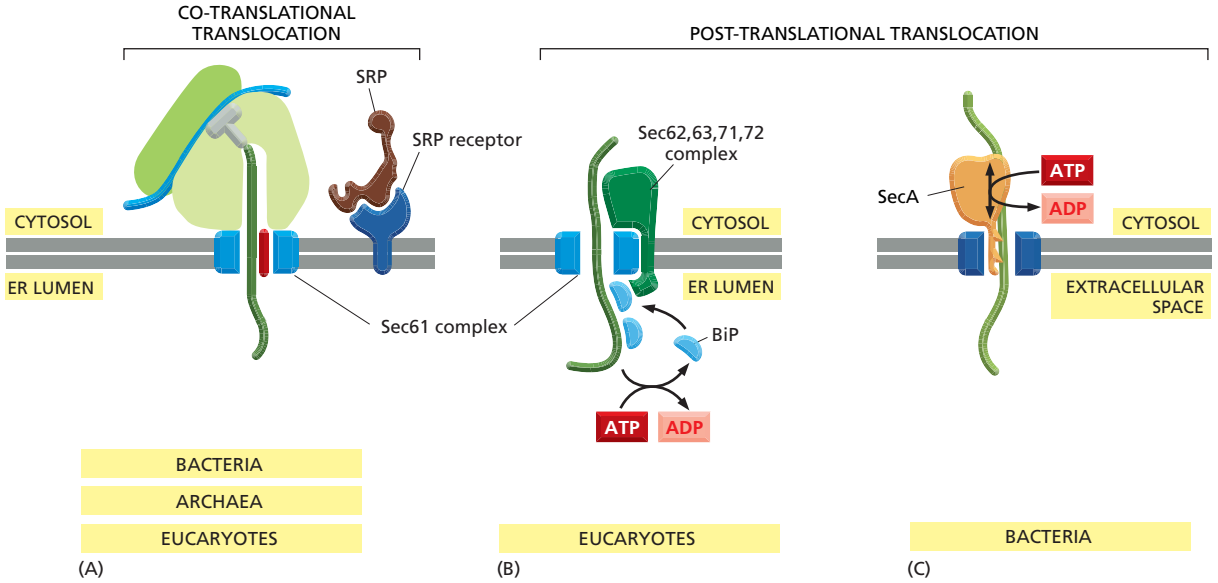


Figure 12–44 Three ways in which protein translocation can be driven through structurally similar translocators. (A) Co-translational translocation. The ribosome is brought to the membrane by the SRP and SRP receptor and engages with the Sec61 protein translocator. The growing polypeptide chain is threaded across the membrane as it is made. No additional energy is needed, as the only path available to the growing chain is to cross the membrane. (B) Post-translational translocation in eucaryotic cells. An additional complex composed of the Sec62, Sec63, Sec71, and Sec72 proteins is attached to the Sec61 translocator and deposits BiP molecules onto the translocating chain as it emerges into the ER lumen. ATP-driven cycles of BiP binding and release pull the protein into the lumen, a mechanism that closely resembles the mechanism of mitochondrial import in Figure 12–26. (C) Post-translational translocation in bacteria. The completed polypeptide chain is fed from the cytosolic side into a translocator in the plasma membrane by the SecA ATPase. ATP-hydrolysis-driven conformational changes drive a pistonlike motion in SecA, each cycle pushing about 20 amino acids of the protein chain through the pore of the translocator. The Sec pathway used for protein translocation across the thylakoid membrane in chloroplasts uses a similar mechanism (see Figure 12–29B).

Whereas the Sec61 translocator, SRP, and SRP receptor are found in all organisms, SecA is found exclusively in bacteria, and the Sec62, Sec63, Sec71, and Sec72 proteins are found exclusively in eucaryotic cells. (Adapted from P. Walter and A.E. Johnson, *Annu. Rev. Cell Biol.* 10:87–119, 1994. With permission from Annual Reviews.)

In Single-Pass Transmembrane Proteins, a Single Internal ER Signal Sequence Remains in the Lipid Bilayer as a Membrane-Spanning α Helix

The signal sequence in the growing polypeptide chain is thought to trigger the opening of the pore in the protein translocator: after the signal sequence is released from the SRP and the growing chain has reached a sufficient length, the signal sequence binds to a specific site inside the pore itself, thereby opening the pore. An ER signal sequence is therefore recognized twice: first by an SRP in the cytosol and then by a binding site in the pore of the protein translocator, where it serves as a **start-transfer signal** (or start-transfer peptide) that opens the pore (illustrated for an ER soluble protein in **Figure 12–45**). Dual recognition may help to ensure that only appropriate proteins enter the lumen of the ER.

While bound in the translocation pore, a signal sequence is in contact not only with the Sec61 complex, which forms the walls of the pore, but also with the hydrophobic lipid core of the membrane. This was shown in chemical cross-linking experiments in which signal sequences and the hydrocarbon chains of lipids could be covalently linked together. When the nascent polypeptide chain grows long enough, an ER signal peptidase cleaves off signal sequences and releases them from the pore into the membrane, where they are rapidly degraded to amino acids by other proteases in the ER membrane. To release the signal sequence into the membrane, the translocator has to open laterally. The translocator is therefore gated in two directions: it can open to form a pore across the membrane to let the hydrophilic portions of proteins cross the lipid bilayer, and it can open laterally within the membrane to let hydrophobic portions of

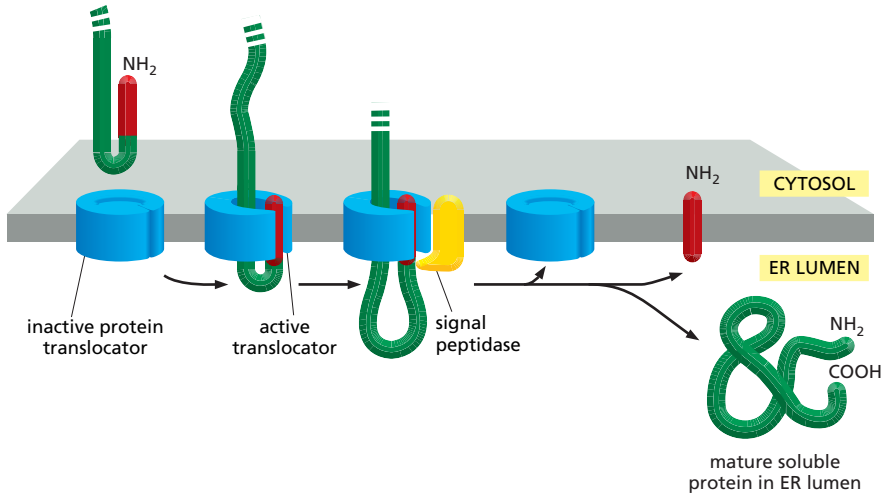


Figure 12–45 A model to explain how a soluble protein is translocated across the ER membrane. On binding an ER signal sequence (which acts as a start-transfer signal), the translocator opens its pore, allowing the transfer of the polypeptide chain across the lipid bilayer as a loop. After the protein has been completely translocated, the pore closes, but the translocator now opens laterally within the lipid bilayer, allowing the hydrophobic signal sequence to diffuse into the bilayer, where it is rapidly degraded. (In this figure and the three figures that follow, the ribosomes have been omitted for clarity.)

proteins partition into the lipid bilayer. Lateral gating of the pore is an essential step during the integration of membrane proteins.

The integration of membrane proteins requires that some parts of the polypeptide chain be translocated across the lipid bilayer whereas others are not. Despite this additional complexity, all modes of insertion of membrane proteins are variants of the sequence of events just described for transferring a soluble protein into the lumen of the ER. We begin by describing the three ways in which **single-pass transmembrane proteins** (see Figure 10–19) become inserted into the ER.

In the simplest case, an N-terminal signal sequence initiates translocation, just as for a soluble protein, but an additional hydrophobic segment in the polypeptide chain stops the transfer process before the entire polypeptide chain is translocated. This **stop-transfer signal** anchors the protein in the membrane after the ER signal sequence (the start-transfer signal) has been released from the translocator and has been cleaved off (**Figure 12–46**). The lateral gating mechanism transfers the stop-transfer sequence into the bilayer, and it remains there as a single α-helical membrane-spanning segment, with the N-terminus of the protein on the lumenal side of the membrane and the C-terminus on the cytosolic side.

In the other two cases, the signal sequence is internal, rather than at the N-terminal end of the protein. Like the N-terminal ER signal sequences, the SRP also

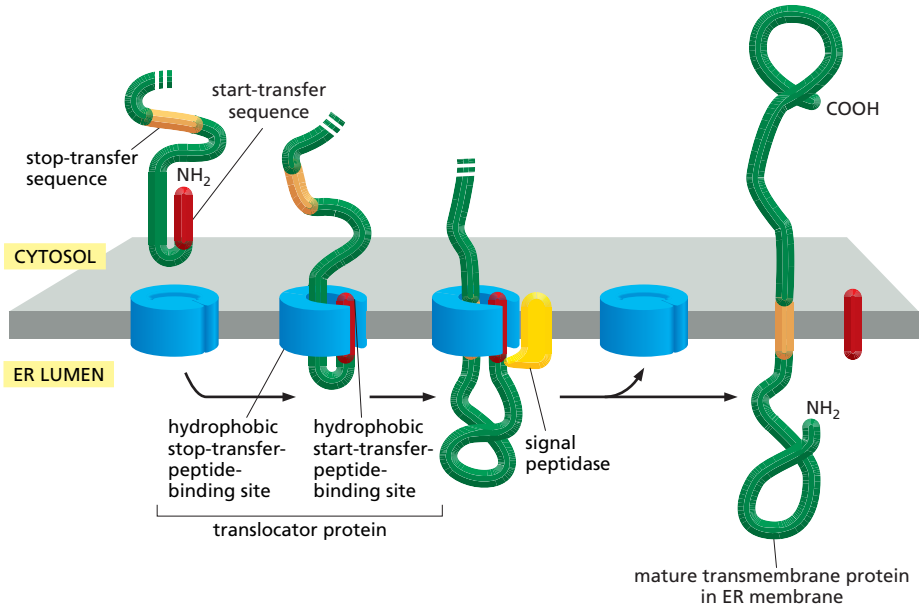


Figure 12–46 How a single-pass transmembrane protein with a cleaved ER signal sequence is integrated into the ER membrane. In this protein the co-translational translocation process is initiated by an N-terminal ER signal sequence *(red)* that functions as a start-transfer signal, as in Figure 12–45. In addition to this start-transfer sequence, however, the protein also contains a stop-transfer sequence *(orange)*. When the stop-transfer sequence enters the translocator and interacts with a binding site, the translocator changes its conformation and discharges the protein laterally into the lipid bilayer.

binds to an internal signal sequence. The SRP brings the ribosome making the protein to the ER membrane and serves as a start-transfer signal that initiates the protein's translocation. After release from the translocator, the internal start-transfer sequence remains in the lipid bilayer as a single membrane-spanning α helix.

Internal start-transfer sequences can bind to the translocation apparatus in either of two orientations; this in turn determines which protein segment (the one preceding or the one following the start-transfer sequence) is moved across the membrane into the ER lumen. In one case, the resulting membrane protein has its C-terminus on the lumenal side (Pathway A in **Figure 12–47**), while in the other, it has its N-terminus on the lumenal side (Pathway B in Figure 12–47). The orientation of the start-transfer sequence depends on the distribution of nearby charged amino acids, as described in the figure legend.

Combinations of Start-Transfer and Stop-Transfer Signals Determine the Topology of Multipass Transmembrane Proteins

In **multipass transmembrane proteins**, the polypeptide chain passes back and forth repeatedly across the lipid bilayer (see Figure 10–19). It is thought that an internal signal sequence serves as a start-transfer signal in these proteins to initiate translocation, which continues until the translocator encounters a stop-transfer sequence. In double-pass transmembrane proteins, for example, the polypeptide can then be released into the bilayer (**Figure 12–48**).

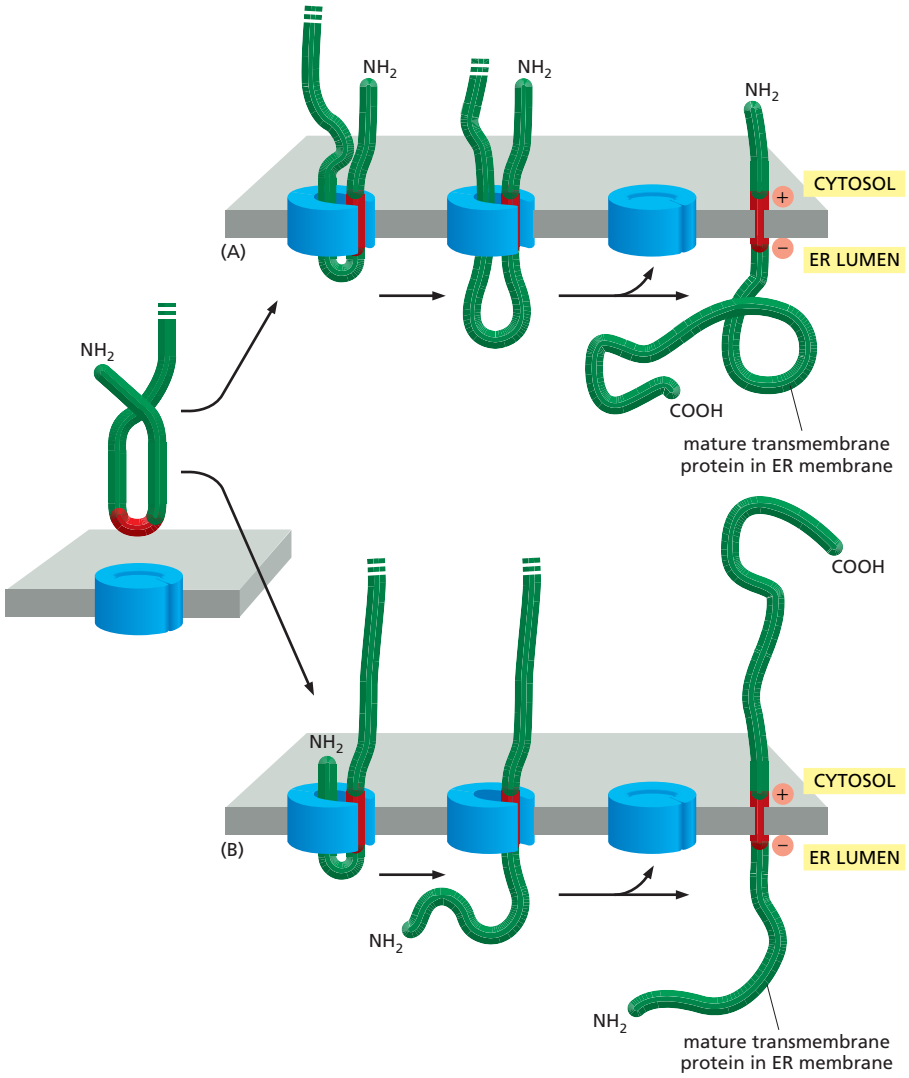


Figure 12–47 Integration of a single-pass transmembrane protein with an internal signal sequence into the ER membrane. An internal ER signal sequence that functions as a start-transfer signal can bind to the translocator in one of two different ways, leading to a membrane protein that has either its C-terminus (pathway A) or its N-terminus (pathway B) in the ER lumen. Proteins are directed into either pathway by features in the polypeptide chain flanking the internal start transfer sequence: if there are more positively charged amino acids immediately *preceding* the hydrophobic core of the start-transfer sequence than there are following it, the membrane protein is inserted into the translocator in the orientation shown in pathway A, whereas if there are more positively charged amino acids immediately *following* the hydrophobic core of the start-transfer sequence than there are preceding it, the membrane protein is inserted into the translocator in the orientation shown in pathway B. Because translocation cannot start before a start-transfer sequence appears outside the ribosome, translocation of the N-terminal portion of the protein shown in (B) can occur only after this portion has been fully synthesized.

Note that there are two ways to insert a single-pass membrane-spanning protein whose N-terminus is located in the ER lumen: that shown in Figure 12–46 and that shown in (B) here.

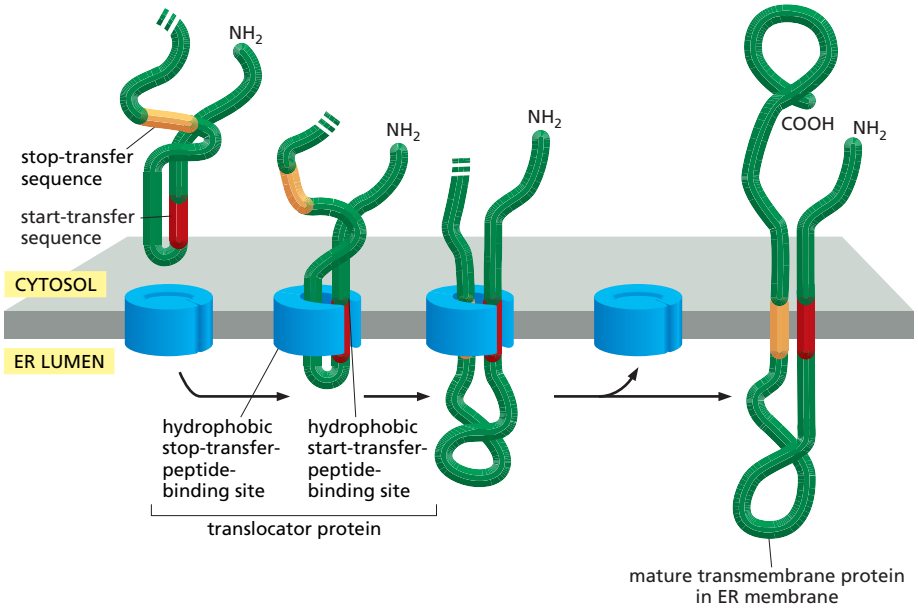


Figure 12–48 Integration of a double-pass transmembrane protein with an internal signal sequence into the ER membrane. In this protein, an internal ER signal sequence acts as a start-transfer signal (as in Figure 12–47) and initiates the transfer of the C-terminal part of the protein. At some point after a stop-transfer sequence has entered the translocator, the translocator discharges the sequence laterally into the membrane.

In more complex multipass proteins, in which many hydrophobic α helices span the bilayer, a second start-transfer sequence reinitiates translocation further down the polypeptide chain until the next stop-transfer sequence causes polypeptide release, and so on for subsequent start-transfer and stop-transfer sequences (**Figure 12–49**).

Whether a given hydrophobic signal sequence functions as a start-transfer or stop-transfer sequence must depend on its location in a polypeptide chain, since its function can be switched by changing its location in the protein by using recombinant DNA techniques. Thus, the distinction between start-transfer and stop-transfer sequences results mostly from their relative order in the growing polypeptide chain. It seems that the SRP begins scanning an unfolded polypeptide chain for hydrophobic segments at its N-terminus and proceeds toward the C-terminus, in the direction that the protein is synthesized. By recognizing the first appropriate hydrophobic segment to emerge from the ribosome, the SRP sets the "reading frame" for membrane integration: after the SRP initiates translocation, the translocator recognizes the next appropriate hydrophobic segment in the direction of transfer as a stop-transfer sequence, causing the region of the polypeptide chain in between to be threaded across the membrane. A similar scanning process continues until all of the hydrophobic regions in the protein have been inserted into the membrane.

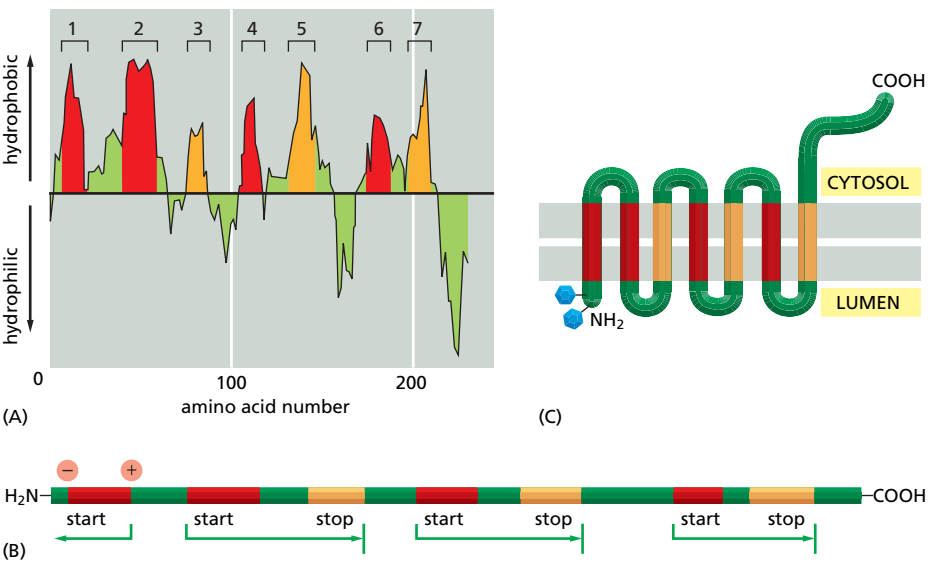


Figure 12–49 The insertion of the multipass membrane protein rhodopsin into the ER membrane. Rhodopsin is the light-sensitive protein in rod photoreceptor cells in the mammalian retina (discussed in Chapter 15). (A) A hydrophobicity plot identifies seven short hydrophobic regions in rhodopsin. (B) The hydrophobic region nearest the N-terminus serves as a start-transfer sequence that causes the preceding N-terminal portion of the protein to pass across the ER membrane. Subsequent hydrophobic sequences function in alternation as start-transfer and stop-transfer sequences. (C) The final integrated rhodopsin has its N-terminus located in the ER lumen and its C-terminus located in the cytosol. The *blue hexagons* represent covalently attached oligosaccharides. Arrows indicate the paired start and stop signals inserted into the translocator.

Because membrane proteins are always inserted from the cytosolic side of the ER in this programmed manner, all copies of the same polypeptide chain will have the same orientation in the lipid bilayer. This generates an asymmetrical ER membrane in which the protein domains exposed on one side are different from those exposed on the other side. This asymmetry is maintained during the many membrane budding and fusion events that transport the proteins made in the ER to other cell membranes (discussed in Chapter 13). Thus, the way in which a newly synthesized protein is inserted into the ER membrane determines the orientation of the protein in all of the other membranes as well.

When proteins dissociate from a membrane and are then reconstituted into artificial lipid vesicles, a random mixture of right-side-out and inside-out protein orientations usually results. Thus, the protein asymmetry observed in cell membranes seems not to be an inherent property of the protein, but instead results solely from the process by which proteins are inserted into the ER membrane from the cytosol.

Translocated Polypeptide Chains Fold and Assemble in the Lumen of the Rough ER

Many of the proteins in the lumen of the ER are in transit, *en route* to other destinations; others, however, normally reside there and are present at high concentrations. These **ER resident proteins** contain an **ER retention signal** of four amino acids at their C-terminus that is responsible for retaining the protein in the ER (see Table 12–3; discussed in Chapter 13). Some of these proteins function as catalysts that help the many proteins that are translocated into the ER to fold and assemble correctly.

One important ER resident protein is *protein disulfide isomerase (PDI)*, which catalyzes the oxidation of free sulfhydryl (SH) groups on cysteines to form disulfide (S–S) bonds. Almost all cysteines in protein domains exposed to either the extracellular space or the lumen of organelles in the secretory and endocytic pathways are disulfide-bonded. By contrast, disulfide bonds form only very rarely in domains exposed to the cytosol because of the reducing environment there.

Another ER resident protein is the chaperone protein **BiP**. We have already discussed how BiP pulls proteins post-translationally into the ER through the ER translocator. Like other chaperones, BiP recognizes incorrectly folded proteins, as well as protein subunits that have not yet assembled into their final oligomeric complexes. It does so by binding to exposed amino acid sequences that would normally be buried in the interior of correctly folded or assembled polypeptide chains. An example of a BiP-binding site is a stretch of alternating hydrophobic and hydrophilic amino acids that would normally be buried in a β sheet. The bound BiP both prevents the protein from aggregating and helps to keep it in the ER (and thus out of the Golgi apparatus and later parts of the secretory pathway). Like some other members of the Hsp70 family of proteins, which bind unfolded proteins and facilitate their import into mitochondria and chloroplasts, BiP hydrolyzes ATP to provide the energy needed to help proteins translocate post-translationally into the ER. It also helps these and other proteins fold.

Most Proteins Synthesized in the Rough ER Are Glycosylated by the Addition of a Common *N*-Linked Oligosaccharide

The covalent addition of sugars to proteins is one of the major biosynthetic functions of the ER. About half of all eucaryotic proteins are glycosylated. Most of the soluble and membrane-bound proteins that are made in the ER—including those destined for transport to the Golgi apparatus, lysosomes, plasma membrane, or extracellular space—are **glycoproteins**. In contrast, very few proteins in the cytosol are glycosylated, and those that are carry a much simpler sugar modification, in which a single *N*-acetylglucosamine group is added to a serine or threonine residue of the protein.

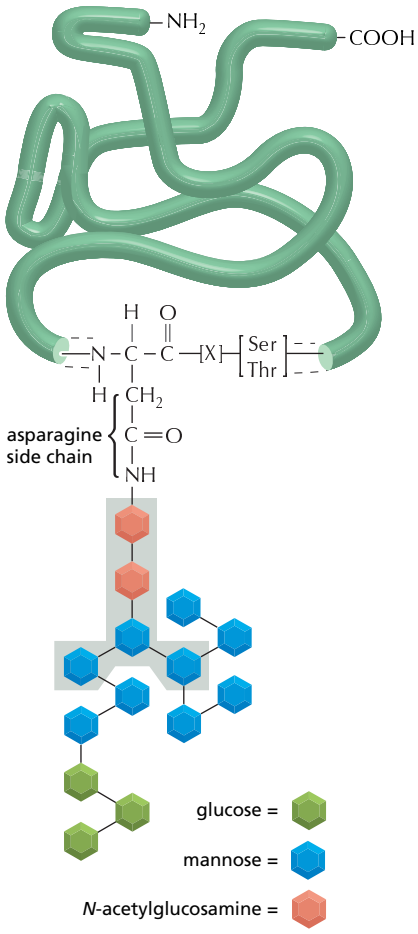


Figure 12–50 The asparagine-linked (*N*-linked) precursor oligosaccharide that is added to most proteins in the rough ER membrane. The five sugars in the *gray box* form the "core region" of this oligosaccharide. For many glycoproteins, only the core sugars survive the extensive oligosaccharide trimming process that takes place in the Golgi apparatus. Only asparagines in the sequences Asn-X-Ser and Asn-X-Thr (where X is any amino acid except proline) become glycosylated. These two sequences occur much less frequently in glycoproteins than in nonglycosylated cytosolic proteins; evidently there has been selective pressure against these sequences during protein evolution, presumably because glycosylation at too many sites would interfere with protein folding.

An important advance in understanding the process of **protein glycosylation** was the discovery that a preformed *precursor oligosaccharide* (composed of *N*-acetylglucosamine, mannose, and glucose and containing a total of 14 sugars) is transferred *en bloc* to proteins in the ER. Because this oligosaccharide is transferred to the side-chain NH_2 group of an asparagine amino acid in the protein, it is said to be *N-linked* or *asparagine-linked* (**Figure 12–50**). The transfer is catalyzed by a membrane-bound enzyme complex, an *oligosaccharyl transferase*, which has its active site exposed on the lumenal side of the ER membrane; this explains why cytosolic proteins are not glycosylated in this way. A special lipid molecule called **dolichol** holds the precursor oligosaccharide in the ER membrane. It transfers its oligosaccharide chain to the target asparagine in a single enzymatic step immediately after that amino acid has reached the ER lumen during protein translocation (**Figure 12–51**). One copy of oligosaccharyl transferase is associated with each protein translocator, allowing it to scan and glycosylate the incoming polypeptide chains efficiently.

The precursor oligosaccharide is linked to the dolichol lipid by a high-energy pyrophosphate bond, which provides the activation energy that drives the glycosylation reaction illustrated in Figure 12–51. The entire precursor oligosaccharide is built up sugar by sugar on this membrane-bound lipid molecule and is then transferred to a protein. The sugars are first activated in the cytosol by the formation of *nucleotide-sugar intermediates*, which then donate their sugar (directly or indirectly) to the lipid in an orderly sequence. Partway

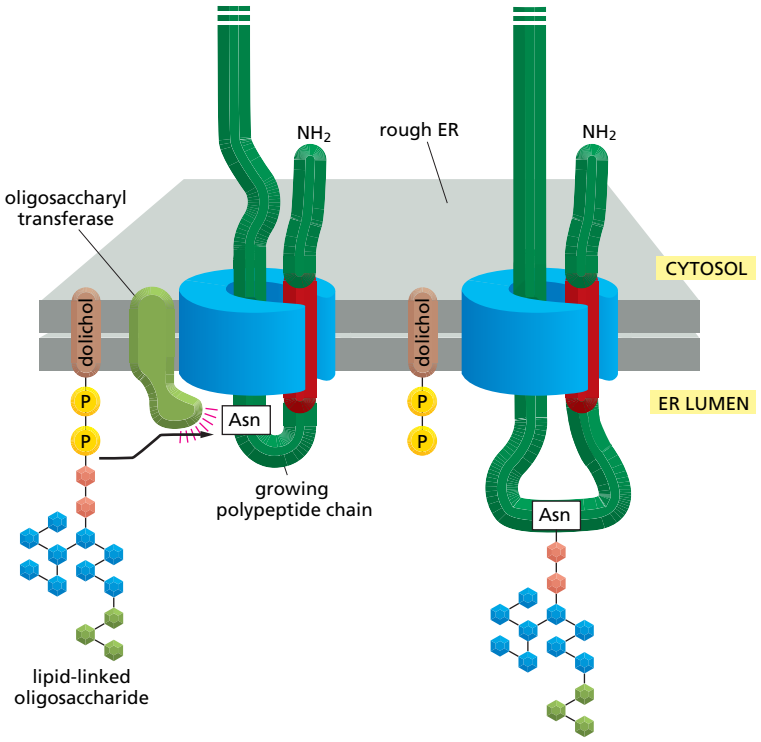


Figure 12–51 Protein glycosylation in the rough ER. Almost as soon as a polypeptide chain enters the ER lumen, it is glycosylated on target asparagine amino acids. The precursor oligosaccharide shown in Figure 12–50 is transferred to the asparagine as an intact unit in a reaction catalyzed by a membrane-bound *oligosaccharyl transferase* enzyme. As with signal peptidase, one copy of this enzyme is associated with each protein translocator in the ER membrane. (The ribosome is not shown.)

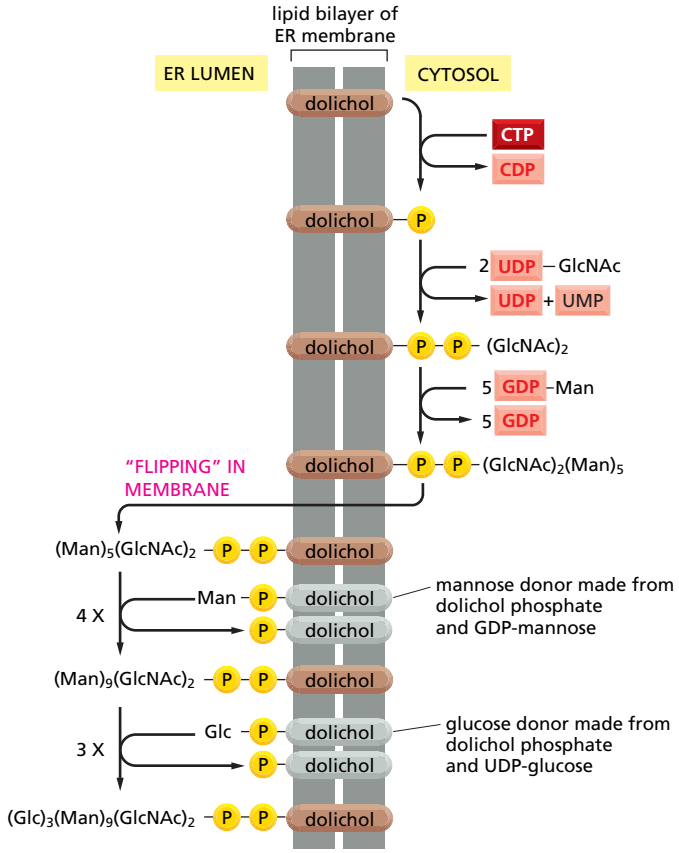


Figure 12–52 Synthesis of the lipid-linked precursor oligosaccharide in the rough ER membrane. The oligosaccharide is assembled sugar by sugar onto the carrier lipid dolichol (a polyisoprenoid; see Panel 2–5, pp. 114–115). Dolichol is long and very hydrophobic: its 22 five-carbon units can span the thickness of a lipid bilayer more than three times, so that the attached oligosaccharide is firmly anchored in the membrane. The first sugar is linked to dolichol by a pyrophosphate bridge. This high-energy bond activates the oligosaccharide for its eventual transfer from the lipid to an asparagine side chain of a nascent polypeptide on the lumenal side of the rough ER. As indicated, the synthesis of the oligosaccharide starts on the cytosolic side of the ER membrane and continues on the lumenal face after the $(Man)_5(GlcNAc)_2$ lipid intermediate is flipped across the bilayer by a transporter. All the subsequent glycosyl transfer reactions on the lumenal side of the ER involve transfers from dolichol-P-glucose and dolichol-P-mannose; these activated, lipid-linked monosaccharides are synthesized from dolichol phosphate and UDP-glucose or GDP-mannose (as appropriate) on the cytosolic side of the ER and are then thought to be flipped across the ER membrane. GlcNAc = *N*-acetylglucosamine; Man = mannose; Glc = glucose.

through this process, the lipid-linked oligosaccharide is flipped, with the help of a transporter, from the cytosolic to the lumenal side of the ER membrane (**Figure 12–52**).

All of the diversity of the *N*-linked oligosaccharide structures on mature glycoproteins results from the later modification of the original precursor oligosaccharide. While still in the ER, three glucoses (see Figure 12–50) and one mannose are quickly removed from the oligosaccharides of most glycoproteins. We shall return to the importance of glucose trimming shortly. This oligosaccharide "trimming" or "processing" continues in the Golgi apparatus and is discussed in Chapter 13.

The *N*-linked oligosaccharides are by far the most common oligosaccharides, being found on 90% of all glycoproteins. Less frequently, oligosaccharides are linked to the hydroxyl group on the side chain of a serine, threonine, or hydroxylysine amino acid. These *O-linked oligosaccharides* are formed in the Golgi apparatus.

Oligosaccharides Are Used as Tags to Mark the State of Protein Folding

It has long been debated why glycosylation is such a common modification of proteins that enter the ER. One particularly puzzling observation has been that some proteins require *N*-linked glycosylation for proper folding in the ER, yet the precise location of the oligosaccharides attached to the protein's surface does not seem to matter. A clue to the role of glycosylation in protein folding came from studies of two ER chaperone proteins, which are called **calnexin** and **calreticulin** because they require Ca^{2+} for their activities. These chaperones are carbohydrate-binding proteins, or *lectins,* which bind to oligosaccharides on incompletely folded proteins and retain them in the ER. Like other chaperones,

they prevent incompletely folded proteins from becoming irreversibly aggregated. Both calnexin and calreticulin also promote the association of incompletely folded proteins with another ER chaperone, which binds to cysteines that have not yet formed disulfide bonds.

Calnexin and calreticulin recognize *N*-linked oligosaccharides that contain a single terminal glucose, and therefore bind proteins only after two of the three glucoses on the precursor oligosaccharide have been removed by ER glucosidases. When the third glucose has been removed, the protein dissociates from its chaperone and can leave the ER.

How, then, do calnexin and calreticulin distinguish properly folded from incompletely folded proteins? The answer lies in yet another ER enzyme, a glucosyl transferase that keeps adding a glucose to those oligosaccharides that have lost their last glucose. It adds the glucose, however, only to oligosaccharides that are attached to unfolded proteins. Thus, an unfolded protein undergoes continuous cycles of glucose trimming (by glucosidase) and glucose addition (by glycosyl transferase), maintaining an affinity for calnexin and calreticulin until it has achieved its fully folded state (**Figure 12–53**).

Improperly Folded Proteins Are Exported from the ER and Degraded in the Cytosol

Despite all the help from chaperones, many protein molecules (more than 80% for some proteins) translocated into the ER fail to achieve their properly folded or oligomeric state. Such proteins are exported from the ER back into the cytosol, where they are degraded. The mechanism of retrotranslocation, also called *dislocation*, is still unknown but is likely to be similar to other post-translational modes of translocation. For example, like translocation into mitochondria or chloroplasts, chaperone proteins are probably necessary to keep the polypeptide chain in an unfolded state prior to and during transport. Similarly, a source of energy is required to provide directionality to the transport and to pull the protein into the cytosol. Finally, a translocator, possibly composed of some of the same components that are used for forward transport into the ER (such as Sec61), is presumably necessary.

Selecting proteins from the ER for degradation is a challenging process. Misfolded proteins or unassembled protein subunits should be degraded, but folding

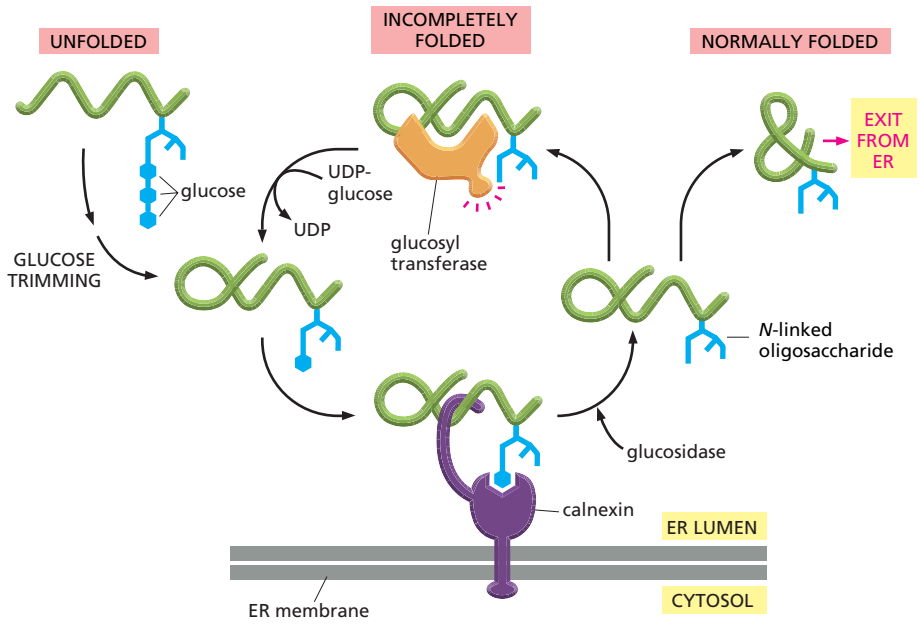


Figure 12–53 The role of *N*-linked glycosylation in ER protein folding. The ER-membrane-bound chaperone protein calnexin binds to incompletely folded proteins containing one terminal glucose on *N*-linked oligosaccharides, trapping the protein in the ER. Removal of the terminal glucose by a glucosidase releases the protein from calnexin. A glucosyl transferase is the crucial enzyme that determines whether the protein is folded properly or not: if the protein is still incompletely folded, the enzyme transfers a new glucose from UDP-glucose to the *N*-linked oligosaccharide, renewing the protein's affinity for calnexin and retaining it in the ER. The cycle repeats until the protein has folded completely. Calreticulin functions similarly, except that it is a soluble ER resident protein. Another ER chaperone, ERp57 (not shown), collaborates with calnexin and calreticulin in retaining an incompletely folded protein in the ER.

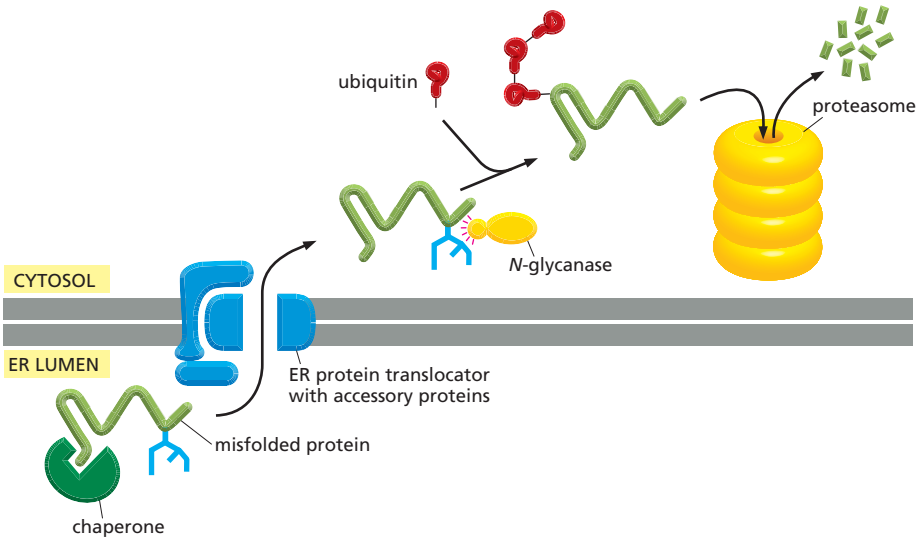


Figure 12–54 The export and degradation of misfolded ER proteins. Misfolded soluble proteins in the ER lumen are translocated back into the cytosol, where they are deglycosylated, ubiquitylated, and degraded in proteasomes. Misfolded membrane proteins follow a similar pathway.

intermediates of newly made proteins should not. Help in making this distinction comes from the *N*-linked oligosaccharides, which serve as timers that measure how long a protein has spent in the ER. The slow trimming of a particular mannose on the core-oligosaccharide tree by an enzyme (a mannosidase) in the ER is thought to create a new oligosaccharide structure that the retrotranslocation apparatus recognizes. Proteins that fold and exit from the ER faster than the action of the mannosidase would therefore escape degradation.

Once a misfolded protein has been retrotranslocated into the cytosol, an *N*-glycanase removes its oligosaccharide chains *en bloc*. The deglycosylated polypeptide is rapidly ubiquitylated by ER-bound ubiquitin-conjugating enzymes and is then fed into proteasomes (discussed in Chapter 6), where it is degraded (**Figure 12–54**).

Misfolded Proteins in the ER Activate an Unfolded Protein Response

Cells carefully monitor the amount of misfolded protein in various compartments. An accumulation of misfolded proteins in the cytosol, for example, triggers a *heat-shock response* (discussed in Chapter 6), which stimulates the transcription of genes encoding cytosolic chaperones that help to refold the proteins. Similarly, an accumulation of misfolded proteins in the ER triggers an **unfolded protein response**, which includes an increased transcription of genes encoding ER chaperones, proteins involved in retrotranslocation and protein degradation in the cytosol, and many other proteins that help to increase the protein folding capacity of the ER.

How do misfolded proteins in the ER signal to the nucleus? There are three parallel pathways that execute the unfolded protein response (**Figure 12–55**A). The first pathway, which was initially discovered in yeast cells, is particularly remarkable. Misfolded proteins in the ER activate a transmembrane protein kinase in the ER, which causes the kinase to oligomerize and phosphorylate itself. (Some cell-surface receptors in the plasma membrane are activated in a similar way, as discussed in Chapter 15). The oligomerization and autophosphorylation activates an endoribonuclease domain in the cytosolic portion of the same molecule, which cleaves a specific, cytosolic RNA molecule at two positions, excising an intron. The separated exons are then joined by an RNA ligase, generating a spliced mRNA, which is translated to produce an active gene regulatory protein. This protein activates the transcription of the genes encoding the proteins that mediate the unfolded protein response (Figure 12–55B).

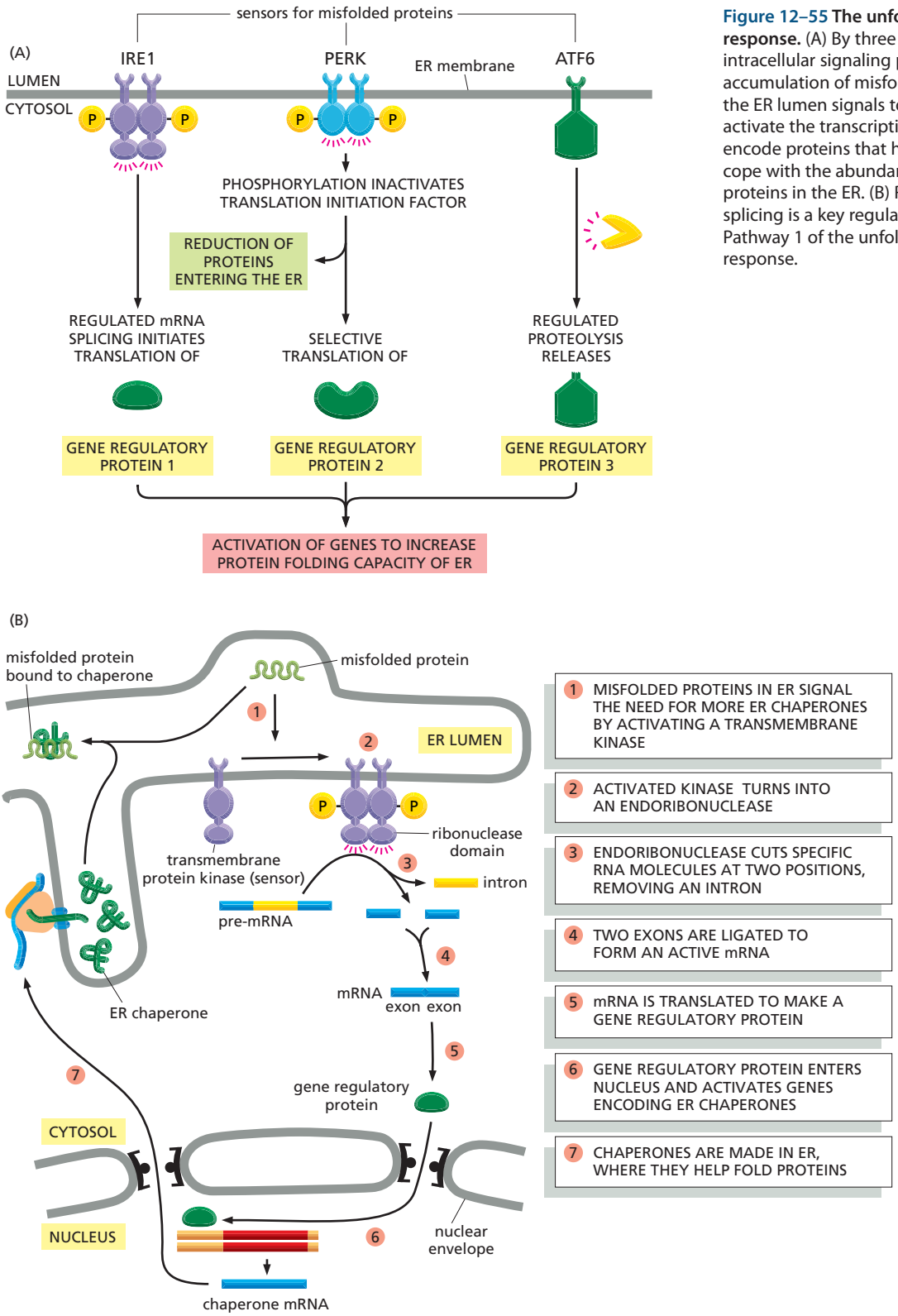


Figure 12–55 The unfolded protein response. (A) By three parallel intracellular signaling pathways, the accumulation of misfolded proteins in the ER lumen signals to the nucleus to activate the transcription of genes that encode proteins that help the cell to cope with the abundance of misfolded proteins in the ER. (B) Regulated mRNA splicing is a key regulatory switch in Pathway 1 of the unfolded protein response.

Misfolded proteins also activate a second transmembrane kinase in the ER, which inhibits a translation initiation factor by phosphorylating it and thereby reduces the production of new proteins throughout the cell. One consequence of the reduction in protein translation is to reduce the flux of proteins into the ER, thereby limiting the load of proteins that need to be folded there. Some proteins, however, are preferentially translated when translation initiation factors

are scarce (see p. 490), and one of these is a gene regulatory protein that helps activate the transcription of the genes encoding proteins active in the unfolded protein response.

Finally, a third gene regulatory protein is initially synthesized as an integral ER membrane protein. Because it is covalently tethered to the membrane, it cannot activate the transcription of genes in the nucleus. When misfolded proteins accumulate in the ER, the transmembrane protein is transported to the Golgi apparatus, where it encounters proteases that cleave off its cytosolic domain, which can now migrate to the nucleus and help activate the transcription of the genes encoding proteins involved in the unfolded protein response. The relative importance of each of these three pathways differs in different cell types, enabling each cell type to tailor the unfolded protein response to its particular needs.

Some Membrane Proteins Acquire a Covalently Attached Glycosylphosphatidylinositol (GPI) Anchor

As discussed in Chapter 10, several cytosolic enzymes catalyze the covalent addition of a single fatty acid chain or prenyl group to selected proteins. The attached lipids help to direct these proteins to cell membranes. A related process is catalyzed by ER enzymes, which covalently attach a **glycosylphosphatidyl-inositol (GPI) anchor** to the C-terminus of some membrane proteins destined for the plasma membrane. This linkage forms in the lumen of the ER, where, at the same time, the transmembrane segment of the protein is cleaved off (**Figure 12–56**). A large number of plasma membrane proteins are modified in this way. Since they are attached to the exterior of the plasma membrane only by their GPI anchors, they can in principle be released from cells in soluble form in response to signals that activate a specific phospholipase in the plasma membrane. Trypanosome

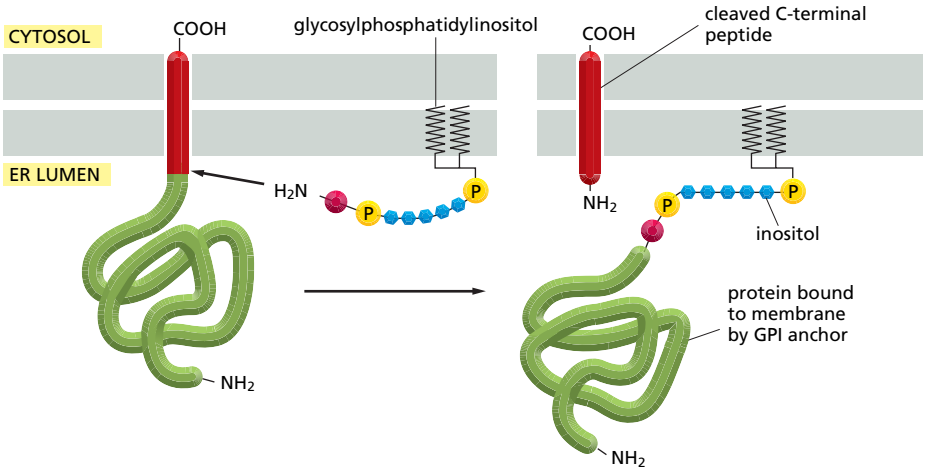


Figure 12–56 The attachment of a GPI anchor to a protein in the ER. Immediately after the completion of protein synthesis, the precursor protein remains anchored in the ER membrane by a hydrophobic C-terminal sequence of 15–20 amino acids; the rest of the protein is in the ER lumen. Within less than a minute, an enzyme in the ER cuts the protein free from its membrane-bound C-terminus and simultaneously attaches the new C-terminus to an amino group on a preassembled GPI intermediate. The sugar chain contains an inositol attached to the lipid from which the GPI anchor derives its name. It is followed by a glucosamine and three mannoses. The terminal mannose links to a phosphoethanolamine that provides the amino group to attach the protein. The signal that specifies this modification is contained within the hydrophobic C-terminal sequence and a few amino acids adjacent to it on the lumenal side of the ER membrane; if this signal is added to other proteins, they too become modified in this way. Because of the covalently linked lipid anchor, the protein remains membrane-bound, with all of its amino acids exposed initially on the lumenal side of the ER and eventually on the cell exterior.

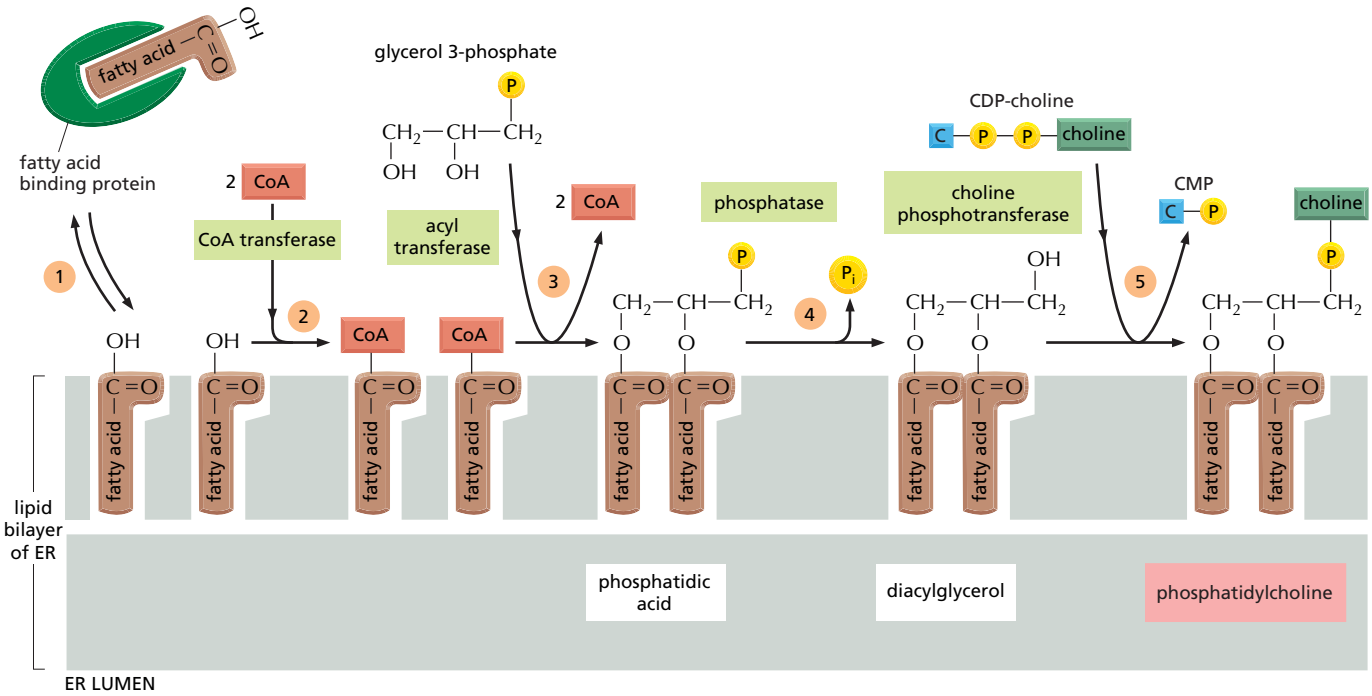


Figure 12–57 The synthesis of phosphatidylcholine. As illustrated, this phospholipid is synthesized from glycerol 3-phosphate, cytidine-diphosphocholine (CDP-choline), and fatty acids delivered to the ER by fatty acid binding protein.

parasites, for example, use this mechanism to shed their coat of GPI-anchored surface proteins when attacked by the immune system. GPI anchors may also be used to direct plasma membrane proteins into *lipid rafts* and thus segregate the proteins from other membrane proteins, as we discuss in Chapter 13.

The ER Assembles Most Lipid Bilayers

The ER membrane synthesizes nearly all of the major classes of lipids, including both phospholipids and cholesterol, required for the production of new cell membranes. The major phospholipid made is *phosphatidylcholine* (also called *lecithin*), which can be formed in three steps from choline, two fatty acids, and glycerol phosphate (**Figure 12–57**). Each step is catalyzed by enzymes in the ER membrane that have their active sites facing the cytosol, where all of the required metabolites are found. Thus, phospholipid synthesis occurs exclusively in the cytosolic leaflet of the ER membrane. Because fatty acids are not soluble in water, they are shepherded from their sites of synthesis to the ER by a fatty acid binding protein in the cytosol. After arrival in the ER membrane and activation with CoA, acyl transferases successively add two fatty acids to glycerol phosphate to produce phosphatidic acid. Phosphatidic acid is sufficiently water-insoluble to remain in the lipid bilayer, and it cannot be extracted from the bilayer by the fatty acid binding proteins. It is therefore this first step that enlarges the ER lipid bilayer. The later steps determine the head group of a newly formed lipid molecule and therefore the chemical nature of the bilayer, but they do not result in net membrane growth. The two other major membrane phospholipids—phosphatidyl ethanolamine and phosphatidylserine—as well as the minor phospholipid phosphatidylinositol (PI), are all synthesized in this way.

Because phospholipid synthesis takes place in the cytosolic half of the ER lipid bilayer, there needs to be a mechanism that transfers some of the newly formed phospholipid molecules to the lumenal leaflet of the bilayer. In synthetic lipid bilayers, lipids do not "flip-flop" in this way. In the ER, however, phospholipids equilibrate across the membrane within minutes, which is almost 100,000

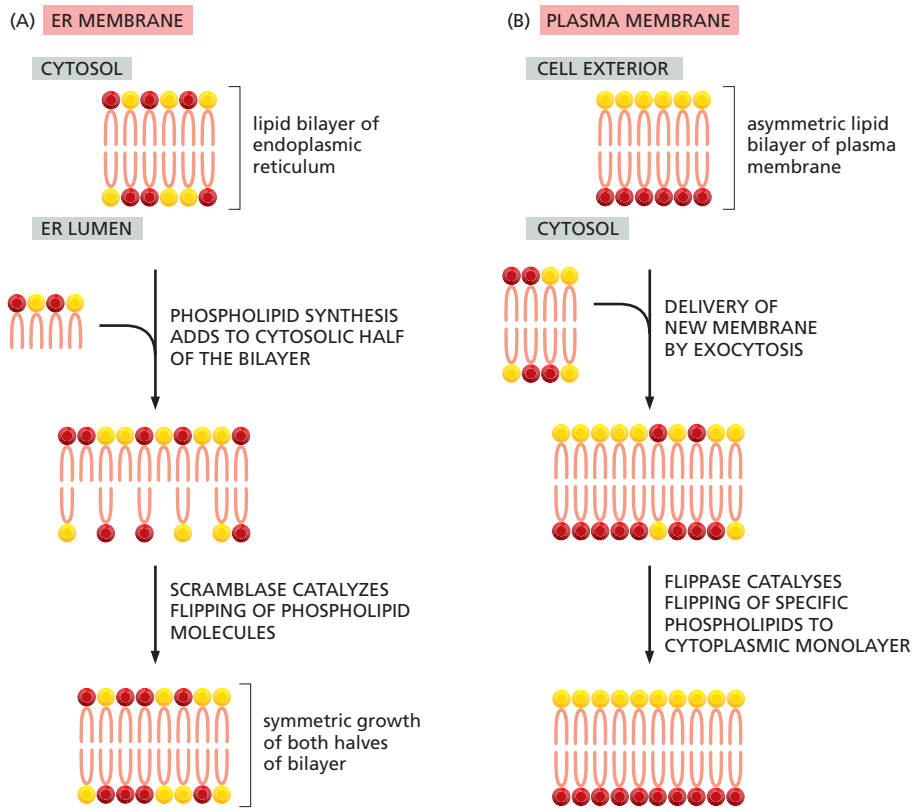


Figure 12–58 The role of phospholipid translocators in lipid bilayer synthesis. (A) Because new lipid molecules are added only to the cytosolic half of the bilayer and lipid molecules do not flip spontaneously from one monolayer to the other, a membrane-bound phospholipid translocator (called a scramblase) is required to transfer lipid molecules from the cytosolic half to the lumenal half so that the membrane grows as a bilayer. The scramblase is not specific for particular phospholipid head groups and therefore equilibrates the different phospholipids between the two monolayers. (B) Fueled by ATP hydrolysis, a head-group-specific flippase in the plasma membrane actively flips phosphatidylserine and phosphatidyl-ethanolamine directionally from the extracellular to the cytosolic leaflet, creating the characteristically asymmetric lipid bilayer of the plasma membrane of animal cells (see Figure 10–16).

times faster than can be accounted for by spontaneous "flip-flop." This rapid trans-bilayer movement is mediated by a poorly characterized phospholipid translocator called a *scramblase*, which equilibrates phospholipids between the two leaflets of the lipid bilayer (**Figure 12–58**). Thus, the different types of phospholipids are thought to be equally distributed between the two leaflets of the ER membrane.

The plasma membrane contains a different type of phospholipid translocator that belongs to the family of P-type pumps (discussed in Chapter 11). These *flippases* specifically remove phospholipids containing in their head groups free amino groups (phosphatidylserine and phosphatidylethanolamine—see Figure 10–3) from the extracellular leaflet and use the energy of ATP hydrolysis to flip them directionally into the leaflet facing the cytosol. The plasma membrane therefore has a highly asymmetric phospholipid composition, which is actively maintained by the flippases (see Figure 10–16). The plasma membrane also contains a scramblase but, in contrast to the ER scramblase, which is always active, the plasma membrane enzyme is regulated and only activated in some situations, such as in apoptosis and in activated platelets, where it acts to abolish lipid asymmetry; the resulting exposure of phosphotidylserine on the surface of apoptotic cells serves as a signal for phagocytic cells to ingest and degrade the dead cell.

The ER also produces cholesterol and ceramide (**Figure 12–59**). *Ceramide* is made by condensing the amino acid serine with a fatty acid to form the amino alcohol *sphingosine* (see Figure 10–3); a second fatty acid is then added to form

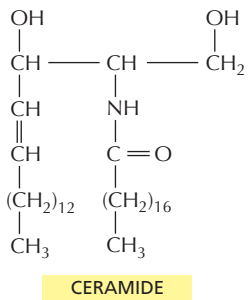


Figure 12–59 The structure of ceramide.

ceramide. The ceramide is exported to the Golgi apparatus, where it serves as a precursor for the synthesis of two types of lipids: oligosaccharide chains are added to form *glycosphingolipids* (glycolipids; see Figure 10–18), and phosphocholine head groups are transferred from phosphatidylcholine to other ceramide molecules to form *sphingomyelin* (discussed in Chapter 10). Thus, both glycolipids and sphingomyelin are produced relatively late in the process of membrane synthesis. Because they are produced by enzymes exposed to the Golgi lumen, they are found exclusively in the noncytosolic leaflet of the lipid bilayers that contain them.

As discussed in Chapter 13, the plasma membrane and the membranes of the Golgi apparatus, lysosomes, and endosomes all form part of a membrane system that communicates with the ER by means of transport vesicles, which transfer both proteins and lipids. Mitochondria and plastids, however, do not belong to this system, and they therefore require different mechanisms to import proteins and lipids for growth. We have already seen that they import most of their proteins from the cytosol. Although mitochondria modify some of the lipids they import, they do not synthesize lipids *de novo;* instead, their lipids have to be imported from the ER, either directly, or indirectly by way of other cell membranes. In either case, special mechanisms are required for the transfer.

The details of how lipid distribution between different membranes is catalyzed and regulated are not known. Water-soluble carrier proteins called *phospholipid exchange proteins* (or *phospholipid transfer proteins)* are thought to transfer individual phospholipid molecules between membranes, functioning much like fatty acid binding proteins that shepherd fatty acids through the cytosol. In addition, mitochondria are often seen in close juxtaposition to ER membranes in electron micrographs, and there may be specific lipid transfer mechanisms that operate between adjacent membranes.

Summary

The extensive ER network serves as a factory for the production of almost all of the cell's lipids. In addition, a major portion of the cell's protein synthesis occurs on the cytosolic surface of the ER: all proteins destined for secretion and all proteins destined for the ER itself, the Golgi apparatus, the lysosomes, the endosomes, and the plasma membrane are first imported into the ER from the cytosol. In the ER lumen, the proteins fold and oligomerize, disulfide bonds are formed, and N-linked oligosaccharides are added. The pattern of N-linked glycosylation is used to indicate the extent of protein folding, so that proteins leave the ER only when they are properly folded. Proteins that do not fold or oligomerize correctly are translocated back into the cytosol, where they are deglycosylated, ubiquitylated, and degraded in proteasomes. If misfolded proteins accumulate in excess in the ER, they trigger an unfolded protein response, which activates appropriate genes in the nucleus to help the ER to cope.

Only proteins that carry a special ER signal sequence are imported into the ER. The signal sequence is recognized by a signal recognition particle (SRP), which binds both the growing polypeptide chain and a ribosome and directs them to a receptor protein on the cytosolic surface of the rough ER membrane. This binding to the ER membrane initiates the translocation process by threading a loop of polypeptide chain across the ER membrane through the hydrophilic pore in a transmembrane protein translocator.

Soluble proteins—destined for the ER lumen, for secretion, or for transfer to the lumen of other organelles—pass completely into the ER lumen. Transmembrane proteins destined for the ER or for other cell membranes are translocated partway across the ER membrane and remain anchored there by one or more membrane-spanning α-helical regions in their polypeptide chains. These hydrophobic portions of the protein can act either as start-transfer or stop-transfer signals during the translocation process. When a polypeptide contains multiple, alternating start-transfer and stop-transfer signals, it will pass back and forth across the bilayer multiple times as a multipass transmembrane protein.

The asymmetry of protein insertion and glycosylation in the ER establishes the sidedness of the membranes of all the other organelles that the ER supplies with membrane proteins.

PROBLEMS

Which statements are true? Explain why or why not.

12–1 Like the lumen of the ER, the interior of the nucleus is topologically equivalent to the outside of the cell.

12–2 Membrane-bound and free ribosomes, which are structurally and functionally identical, differ only in the proteins they happen to be making at a particular time.

12–3 To avoid the inevitable congestion that would occur if two-way traffic through a single pore were allowed, nuclear pore complexes are specialized so that some mediate import while others mediate export.

12–4 Peroxisomes are found in only a few specialized types of eucaryotic cell.

12–5 In multipass transmembrane proteins the odd-numbered transmembrane segments (counting from the N-terminus) act as start-transfer signals and the even-numbered segments act as stop-transfer signals.

Discuss the following problems.

12–6 What is the fate of a protein with no sorting signal?

12–7 Imagine that you have engineered a set of genes, each encoding a protein with a pair of conflicting signal sequences that specify different compartments. If the genes were expressed in a cell, predict which signal would win out for the following combinations. Explain your reasoning.

A. Signals for import into nucleus and import into ER.
B. Signals for import into peroxisomes and import into ER.
C. Signals for import into mitochondria and retention in ER.
D. Signals for import into nucleus and export from nucleus.

12–8 The rough ER is the site of synthesis of many different classes of membrane protein. Some of these proteins remain in the ER, whereas others are sorted to compartments such as the Golgi apparatus, lysosomes, and the plasma membrane. One measure of the difficulty of the sorting problem is the degree of "purification" that must be achieved during transport from the ER. Are proteins bound for the plasma membrane common or rare among all ER membrane proteins?

A few simple considerations allow one to answer this question. In a typical growing cell that is dividing once every 24 hours, the equivalent of one new plasma membrane must transit the ER every day. If the ER membrane is 20 times the area of a plasma membrane, what is the ratio of plasma membrane proteins to other membrane proteins in the ER? (Assume that all proteins on their way to the plasma membrane remain in the ER for 30 minutes on average before exiting, and that the ratio of proteins to lipids in the ER and plasma membranes is the same.)

12–9 Before nuclear pore complexes were well understood, it was unclear whether nuclear proteins diffused passively into the nucleus and accumulated there by binding to residents of the nucleus such as chromosomes, or were actively transported and accumulated regardless of their affinity for nuclear components.

A classic experiment that addressed this problem used several forms of radioactive nucleoplasmin, which is a large pentameric protein involved in chromatin assembly. In this experiment, either the intact protein or the nucleoplasmin heads, tails, or heads with a single tail were injected into the cytoplasm of a frog oocyte or into the nucleus (**Figure Q12–1**). All forms of nucleoplasmin, except heads, accumulated in the nucleus when injected into the cytoplasm, and all forms were retained in the nucleus when injected there.

A. What portion of the nucleoplasmin molecule is responsible for localization in the nucleus?
B. How do these experiments distinguish between active transport, in which a nuclear localization signal triggers transport by the nuclear pore complex, and passive diffusion, in which a binding site for a nuclear component allows accumulation in the nucleus?

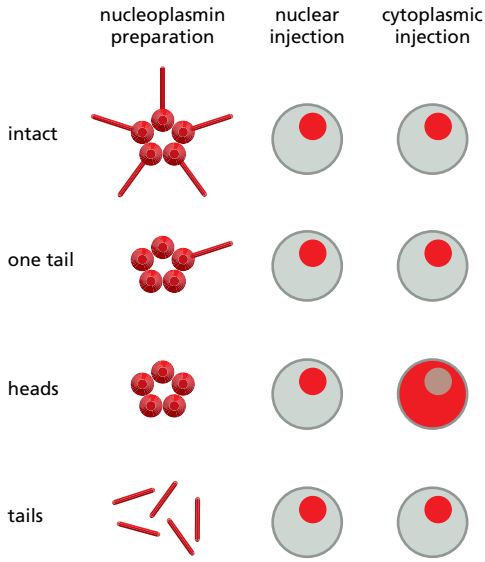


Figure Q12–1 Cell location of injected nucleoplasmin and nucleoplasmin components (Problem 12–9). Schematic diagrams of autoradiographs show the cytoplasm and nucleus with the location of nucleoplasmin indicated by the *red* areas.

12–10 Assuming that 32 million histone octamers are required to package the human genome, how many histone molecules must be transported per second per pore complex in cells whose nuclei contain 3000 nuclear pores and are dividing once per day?

12–11 The structures of Ran-GDP and Ran-GTP (actually Ran-GppNp, a stable GTP analog) are strikingly different, as shown in **Figure Q12–2.** Not surprisingly, Ran-GDP binds to a different set of proteins than does Ran-GTP.

To look at the uptake of Ran itself into the nuclei of permeabilized cells, you attach a red fluorescent tag to a cysteine side chain in Ran to make it visible. This modified Ran supports normal nuclear uptake. Fluorescent Ran-GDP is taken up by nuclei only if cytoplasm is added, whereas a

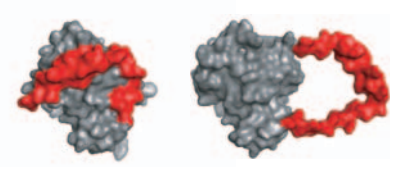


Figure Q12–2 Structures of Ran-GDP and Ran-GppNp (Problem 12–11). The *red portions* of the structures show the segment of Ran that differs most dramatically when GDP or the GTP analog is bound.

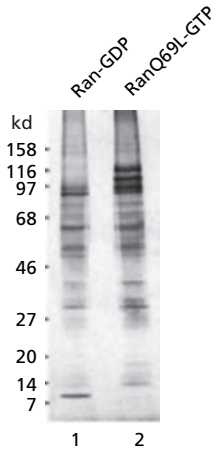


Figure Q12–3 Proteins eluted from Ran-GDP (lane 1) and RanQ69L-GTP (lane 2) affinity columns (Problem 12–11). The molecular weights of marker proteins are shown on the *left*. (From K. Ribbeck et al., *EMBO J.* 17:6587–6598, 1998. With permission from Macmillan Publishers Ltd.)

mutant form, RanQ69L-GTP, which is unable to hydrolyze GTP, is not taken up in the presence or absence of cytoplasm. To identify the cytoplasmic protein that is crucial for Ran-GDP uptake, you construct affinity columns with bound Ran-GDP or bound RanQ69L-GTP and pass cytoplasm through them. Cytoplasm passed over a Ran-GDP column no longer supports nuclear uptake of Ran-GDP, whereas cytoplasm passed over a column of RanQ69L-GTP retains this activity. You elute the bound proteins from each column and analyze them on an SDS polyacrylamide gel, looking for differences that might identify the factor that is required for nuclear uptake of Ran (**Figure Q12–3**).

A. Why did you use RanQ69L-GTP instead of Ran-GTP in these experiments? Would Ran-GppNp in place of RanQ69L-GTP have achieved the same purpose?

B. Which of the many proteins eluted from the two different affinity columns is a likely candidate for the factor that promotes nuclear import of Ran-GDP?

C. What other protein or proteins would you predict that the Ran-GDP import factor would be likely to bind in order to carry out its function?

D. How might you confirm that the factor you have identified is necessary for promoting the nuclear uptake of Ran?

12–12 Components of the TIM complex, the multisubunit protein translocator in the mitochondrial inner membrane, are much less abundant than those of the TOM complex. They were initially identified by using a genetic trick. The yeast *Ura3* gene, whose product is normally located in the cytosol where it is essential for synthesis of uracil, was modified to carry an import signal for the mitochondrial matrix. A population of cells carrying the modified *Ura3* gene was then grown in the absence of uracil. Most cells died, but the rare cells that grew were shown to be defective for mitochondrial import. Explain how this selection identifies cells with defects in components required for mitochondrial import. Why do normal cells with the modified *Ura3* gene not grow in the absence of uracil? Why do cells that are defective for mitochondrial import grow in the absence of uracil?

12–13 If the enzyme dihydrofolate reductase (DHFR), which is normally located in the cytosol, is engineered to carry a mitochondrial targeting sequence at its N-terminus, it is efficiently imported into mitochondria. If the modified DHFR is first incubated with methotrexate, which binds tightly to the active site, the enzyme remains in the cytosol. How do you suppose that the binding of methotrexate interferes with mitochondrial import?

12–14 Why do mitochondria need a fancy translocator to import proteins across the outer membrane? Their outer membranes already have large pores formed by porins.

12–15 Catalase, an enzyme normally found in peroxisomes, is present in normal amounts in cells that do not have visible peroxisomes. It is possible to determine the location of catalase in such cells using immunofluorescence microscopy with antibodies specific for catalase. Fluorescence micrographs of normal cells and peroxisome-deficient cells are shown in **Figure Q12–4**. Where is catalase located in cells without peroxisomes (Figure Q12–4B)? Why does catalase show up as small dots of fluorescence in normal cells (Figure Q12–4A)?

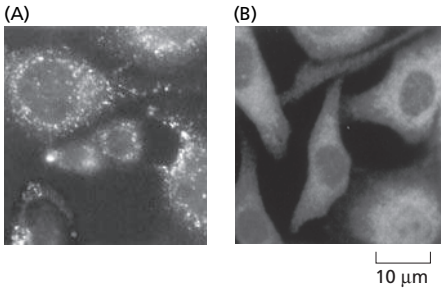


Figure Q12–4 Location of catalase in cells as determined by immunofluorescence microscopy (Problem 12–15). (A) Normal cells. (B) Peroxisome-deficient cells. Cells were reacted with antibodies specific for catalase, washed, and then stained with a fluorescein-labeled second antibody that is specific for the catalase-specific antibody. The two panels are at the same magnification. (From N. Kinoshita et al., *J. Biol. Chem.* 273:24122–24130, 1998. With permission from the American Society for Biochemistry and Molecular Biology.)

12–16 Examine the multipass transmembrane protein shown in **Figure Q12–5**. What would you predict would be the effect of converting the first hydrophobic transmembrane segment to a hydrophilic segment? Sketch the arrangement of the modified protein in the ER membrane.

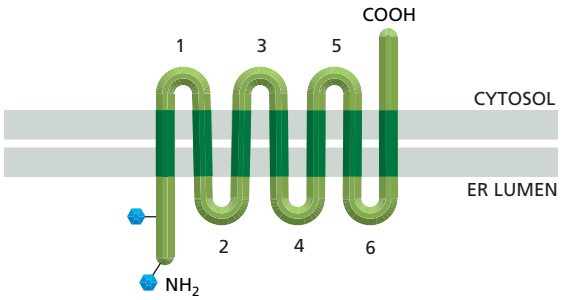


Figure Q12–5 Arrangement of a multipass transmembrane protein in the ER membrane (Problem 12–16). *Hexagons* represent covalently attached oligosaccharides.

12–17 All new phospholipids are added to the cytoplasmic leaflet of the ER membrane, yet the ER membrane has a symmetrical distribution of different phospholipids in its two leaflets. By contrast, the plasma membrane, which receives all its membrane components ultimately from the ER, has a very asymmetrical distribution of phospholipids in the two leaflets of its lipid bilayer. How is symmetry generated in the ER membrane, and how is asymmetry generated in the plasma membrane?

REFERENCES

General

Palade G (1975) Intracellular aspects of the process of protein synthesis. *Science* 189:347–358.

The Compartmentalization of Cells

Blobel G (1980) Intracellular protein topogenesis. *Proc Natl Acad Sci USA* 77:1496–1500.

De Duve C (2007) The origin of eukaryotes: a reappraisal. *Nature Rev Genet* 8:395–403.

Schatz G & Dobberstein B (1996) Common principles of protein translocation across membranes. *Science* 271:1519–1526.

Warren G & Wickner W (1996) Organelle inheritance. *Cell* 84:395–400.

The Transport of Molecules Between the Nucleus and the Cytosol

Adam SA & Gerace L (1991) Cytosolic proteins that specifically bind nuclear location signals are receptors for nuclear import. *Cell* 66:837–847.

Bednenko J, Cingolani G & Gerace L (2003) Nucleocytoplasmic transport: navigating the channel. *Traffic* 4:127–135.

Chook YM & Blobel G (2001) Karyopherins and nuclear import. *Curr Opin Struct Biol* 11:703–715.

Cole CN & Scarcelli JJ (2006) Transport of messenger RNA from the nucleus to the cytoplasm. *Curr Opin Cell Biol* 18:299–306.

Fahrenkrog B, Koser J & Aebi U (2004) The nuclear pore complex: a jack of all trades? *Trends Biochem Sci* 29:175–182.

Görlich D & Kutay U (1999) Transport between the cell nucleus and the cytoplasm. *Annu Rev Cell Dev Biol* 15:607–660.

Hetzer MW, Walther TC & Mattaj IW (2005) Pushing the envelope: structure, function, and dynamics of the nuclear periphery. *Annu Rev Cell Dev Biol* 21:347–380.

Komeili A & O'Shea EK (2000) Nuclear transport and transcription. *Curr Opin Cell Biol* 12:355–360.

Kuersten S, Ohno M & Mattaj IW (2001) Nucleocytoplasmic transport: Ran, beta and beyond. *Trends Cell Biol* 11:497–503.

Tran EJ & Wente SR (2006) Dynamic nuclear pore complexes: life on the edge. *Cell* 125:1041–1053.

Weis K (2002) Nucleocytoplasmic transport: cargo trafficking across the border. *Curr Opin Cell Biol* 14:328–335.

The Transport of Proteins Into Mitochondria and Chloroplasts

Jarvis P & Robinson C (2004) Mechanisms of protein import and routing in chloroplasts. *Curr Biol* 14:R1064–R1077.

Kessler F & Schnell DJ (2004) Chloroplast protein import: solve the GTPase riddle for entry. *Trends Cell Biol* 14:334–338.

Koehler CM, Merchant S & Schatz G (1999) How membrane proteins travel across the mitochondrial intermembrane space. *Trends Biochem Sci* 24:428–432.

Mokranjac D & Neupert W (2005) Protein import into mitochondria. *Biochem Soc Trans* 33:1019–1023.

Prakash S & Matouschek A (2004) Protein unfolding in the cell. *Trends Biochem Sci* 29:593–600.

Soll J & Schleiff E (2004) Protein import into chloroplasts. *Nature Rev Mol Cell Biol* 5:198–208.

Truscott KN, Brandner K & Pfanner N (2003) Mechanisms of protein import into mitochondria. *Curr Biol* 13:R326–R337.

Peroxisomes

Fujiki Y (2000) Peroxisome biogenesis and peroxisome biogenesis disorders. *FEBS Lett* 476:42–46.

Lazarow PB (2003) Peroxisome biogenesis: advances and conundrums. *Curr Opin Cell Biol* 15:489–497.

van der Zand A, Braakman I & Tabak HF (2006) The return of the peroxisome. *J Cell Sci* 119:989–994.

The Endoplasmic Reticulum

Adelman MR, Sabatini DD et al (1973) Ribosome-membrane interaction. Nondestructive disassembly of rat liver rough microsomes into ribosomal and membranous components. *J Cell Biol* 56:206–229.

Bernales S, Papa FR & Walter P (2006) Intracellular signaling by the unfolded protein response. *Annu Rev Cell Dev Biol* 22:487–508.

Bishop WR & Bell RM (1988) Assembly of phospholipids into cellular membranes: biosynthesis, transmembrane movement and intracellular translocation. *Annu Rev Cell Biol* 4:579–610.

Blobel G & Dobberstein B (1975) Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J Cell Biol* 67:835–851.

Borgese N, Mok W & Sabatini DD (1974) Ribosomal-membrane interaction: in vitro binding of ribosomes to microsomal membranes. *J Mol Biol* 88:559–580.

Daleke DL (2003) Regulation of transbilayer plasma membrane phospholipid asymmetry. *J Lipid Res* 44:233–242.

Deshaies RJ, Sanders SL & Schekman R (1991) Assembly of yeast Sec proteins involved in translocation into the endoplasmic reticulum into a membrane-bound multisubunit complex. *Nature* 349:806–808.

Ellgaard L & Helenius A (2003) Quality control in the endoplasmic reticulum. *Nature Rev Mol Cell Biol* 4:181–191.

Ferguson MA (1999) The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research. *J Cell Sci* 112:2799–2809.

Gething MJ (1999) Role and regulation of the ER chaperone BiP. *Semin Cell Dev Biol* 10:465–472.

Görlich D, Prehn S et al (1992) A mammalian homolog of SEC61p and SECYp is associated with ribosomes and nascent polypeptides during translocation. *Cell* 71:489–503.

Helenius J & Aebi M (2002) Transmembrane movement of dolichol linked carbohydrates during N-glycoprotein biosynthesis in the endoplasmic reticulum. *Semin Cell Dev Biol* 13:171–178.

Johnson AE & van Waes MA (1999) The translocon: a dynamic gateway at the ER membrane. *Annu Rev Cell Dev Biol* 15:799–842.

Keenan RJ, Freymann DM & Walter P (2001) The signal recognition particle. *Annu Rev Biochem* 70:755–775.

Kostova Z & Wolf DH (2003) For whom the bell tolls: protein quality control of the endoplasmic reticulum and the ubiquitin-proteasome connection. *EMBO J* 22:2309–2317.

Levine T & Loewen C (2006) Inter-organelle membrane contact sites: through a glass, darkly. *Curr Opin Cell Biol* 18:371–378.

Marciniak SJ & Ron D (2006) Endoplasmic reticulum stress signaling in disease. *Physiol Rev* 86:1133–1149.

Milstein C, Brownlee GG et al (1972) A possible precursor of immunoglobulin light chains. *Nature New Biol* 239:117–120.

Römisch K (2005) Endoplasmic reticulum-associated degradation. *Annu Rev Cell Dev Biol* 21:435–456.

Simon SM & Blobel G (1991) A protein-conducting channel in the endoplasmic reticulum. *Cell* 65:371–380.

Staehelin LA (1997) The plant ER: a dynamic organelle composed of a large number of discrete functional domains. *Plant J* 11:1151–1165.

Trombetta ES & Parodi AJ (2003) Quality control and protein folding in the secretory pathway. *Annu Rev Cell Dev Biol* 19:649–676.

Tsai B, Ye Y & Rapoport TA (2002) Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nature Rev Mol Cell Biol* 3:246–255.

Shibata Y, Voeltz GK & Rapoport TA (2006) Rough sheets and smooth tubules. *Cell* 126:435–439

White SH & von Heijne G (2004) The machinery of membrane protein assembly. *Curr Opin Struct Biol* 14:397–404.

Yan A & Lennarz WJ (2005) Unraveling the mechanism of protein N-glycosylation. *J Biol Chem* 280:3121–3124.

13 Intracellular Vesicular Traffic

Every cell must eat, communicate with the world around it, and quickly respond to changes in its environment. To help accomplish these tasks, cells continually adjust the composition of their plasma membrane in rapid response to need. They use an elaborate internal membrane system to add and remove cell-surface proteins embedded in the membrane, such as receptors, ion channels, and transporters. Through the process of *exocytosis*, the *biosynthetic–secretory pathway* delivers newly synthesized proteins, carbohydrates, and lipids to either the plasma membrane or the extracellular space. By the converse process of *endocytosis* (**Figure 13–1**) cells remove plasma membrane components and deliver them to internal compartments called *endosomes*, from where they can be recycled to the same or different regions of the plasma membrane or can be delivered to lysosomes for degradation. Cells also use endocytosis to capture important nutrients, such as vitamins, lipids, cholesterol, and iron; these are taken up together with the macromolecules to which they bind and are then released in endosomes or lysosomes and transported into the cytosol, where they are used in various biosynthetic processes.

The interior space, or *lumen*, of each membrane-enclosed compartment along the biosynthetic–secretory and endocytic pathways is topologically equivalent to the lumen of most other membrane-enclosed compartments and to the cell exterior. Proteins can travel in this space without having to cross a membrane, being passed from one compartment to another by means of numerous membrane-enclosed transport containers. Some of these containers are small spherical *vesicles*, while others are larger irregular vesicles or tubules formed from the donor compartment. We shall use the term **transport vesicle** to apply to all forms of these containers.

Within a eucaryotic cell, transport vesicles continually bud off from one membrane and fuse with another, carrying membrane components and soluble molecules, which are referred to as **cargo** (**Figure 13–2**). This membrane traffic flows along highly organized, directional routes, which allow the cell to secrete, eat, and remodel its plasma membrane. The biosynthetic–secretory pathway leads outward from the endoplasmic reticulum (ER) toward the Golgi apparatus and cell surface, with a side route leading to lysosomes, while the endocytic pathway leads inward from the plasma membrane. In each case, the flow of membrane between compartments is balanced, with retrieval pathways balancing the flow in the opposite direction, bringing membrane and selected proteins back to the compartment of origin (**Figure 13–3**).

To perform its function, each transport vesicle that buds from a compartment must be selective. It must take up only the appropriate molecules and must fuse only with the appropriate target membrane. A vesicle carrying cargo from the Golgi apparatus to the plasma membrane, for example, must exclude proteins that are to stay in the Golgi apparatus, and it must fuse only with the plasma membrane and not with any other organelle.

We begin this chapter by considering the molecular mechanisms of budding and fusion that underlie all vesicular transport. We then discuss the fundamental problem of how, in the face of this transport, the cell maintains the differences between the compartments. Finally, we consider the function of the Golgi apparatus, lysosomes, secretory vesicles, and endosomes, as we trace the pathways that connect these organelles.

In This Chapter

THE MOLECULAR MECHANISMS OF MEMBRANE TRANSPORT AND THE MAINTENANCE OF COMPARTMENT DIVERSITY 750

TRANSPORT FROM THE ER THROUGH THE GOLGI APPARATUS 766

TRANSPORT FROM THE *TRANS* GOLGI NETWORK TO LYSOSOMES 779

TRANSPORT INTO THE CELL FROM THE PLASMA MEMBRANE: ENDOCYTOSIS 787

TRANSPORT FROM THE *TRANS* GOLGI NETWORK TO THE CELL EXTERIOR: EXOCYTOSIS 799

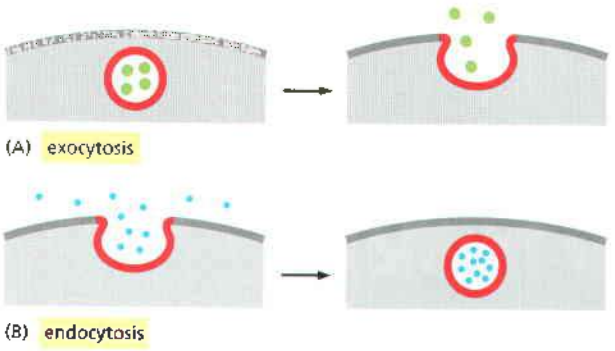


Figure 13–1 **Exocytosis and endocytosis.** (A) In exocytosis, a transport vesicle fuses with the plasma membrane. Its content is released into the extracellular space, while the vesicle membrane (*red*) becomes continuous with the plasma membrane. (B) In endocytosis, a plasma membrane patch (*red*) is internalized forming a transport vesicle. Its content derives from the extracellular space.

THE MOLECULAR MECHANISMS OF MEMBRANE TRANSPORT AND THE MAINTENANCE OF COMPARTMENTAL DIVERSITY

Vesicular transport mediates a continuous exchange of components between the ten or more chemically distinct, membrane-enclosed compartments that collectively comprise the biosynthetic-secretory and endocytic pathways. With this massive exchange, how can each compartment maintain its special identity? To answer this question, we must first consider what defines the character of a compartment. Above all, it is the composition of the enclosing membrane: molecular markers displayed on the cytosolic surface of the membrane serve as guidance cues for incoming traffic to ensure that transport vesicles fuse only with the correct compartment. Many of these membrane markers, however, are found on more than one compartment, and it is the specific combination of marker molecules that gives each compartment its unique molecular address.

How are these membrane markers kept at high concentration on one compartment and at low concentration on another? To answer this question, we need to consider how patches of membrane, enriched or depleted in specific membrane components, bud off from one compartment and transfer to another. Panel 13–1 outlines some of the genetic and biochemical strategies that have been used to study the molecular machinery involved in vesicular transport.

We begin by discussing how cells segregate proteins into separate membrane domains by assembling a special protein coat on the membrane's cytosolic face. We consider how coats form, what they are made of, and how they are

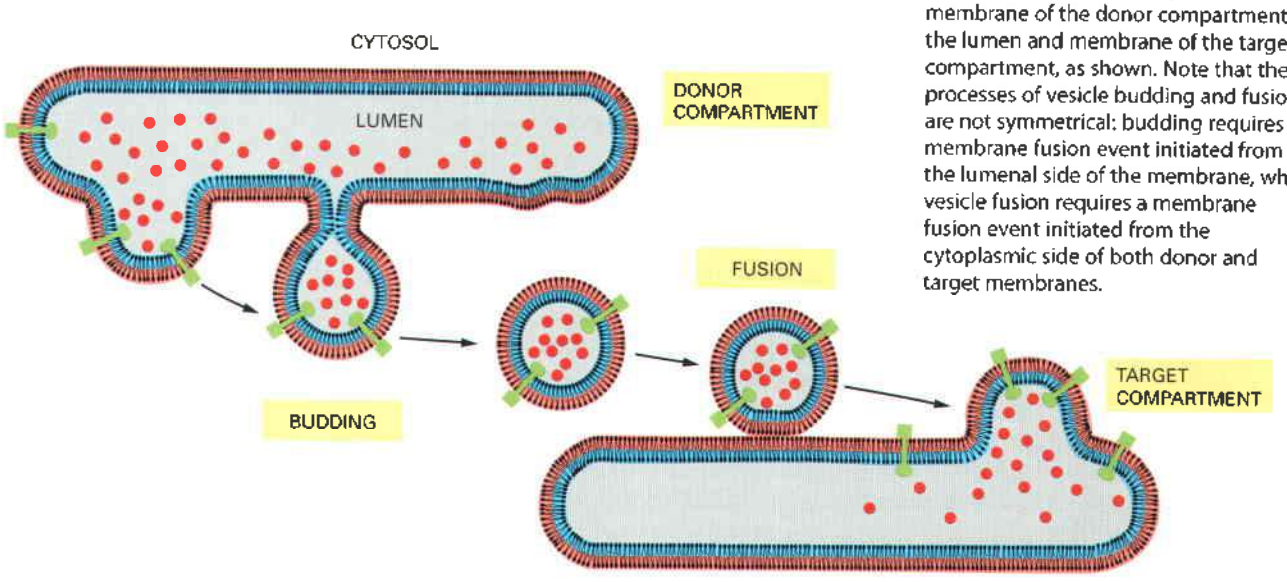


Figure 13–2 **Vesicular transport.** Transport vesicles bud off from one compartment and fuse with another. As they do so, they carry material as cargo from the *lumen* (the space within a membrane-enclosed compartment) and membrane of the donor compartment to the lumen and membrane of the target compartment, as shown. Note that the processes of vesicle budding and fusion are not symmetrical: budding requires a membrane fusion event initiated from the lumenal side of the membrane, while vesicle fusion requires a membrane fusion event initiated from the cytoplasmic side of both donor and target membranes.

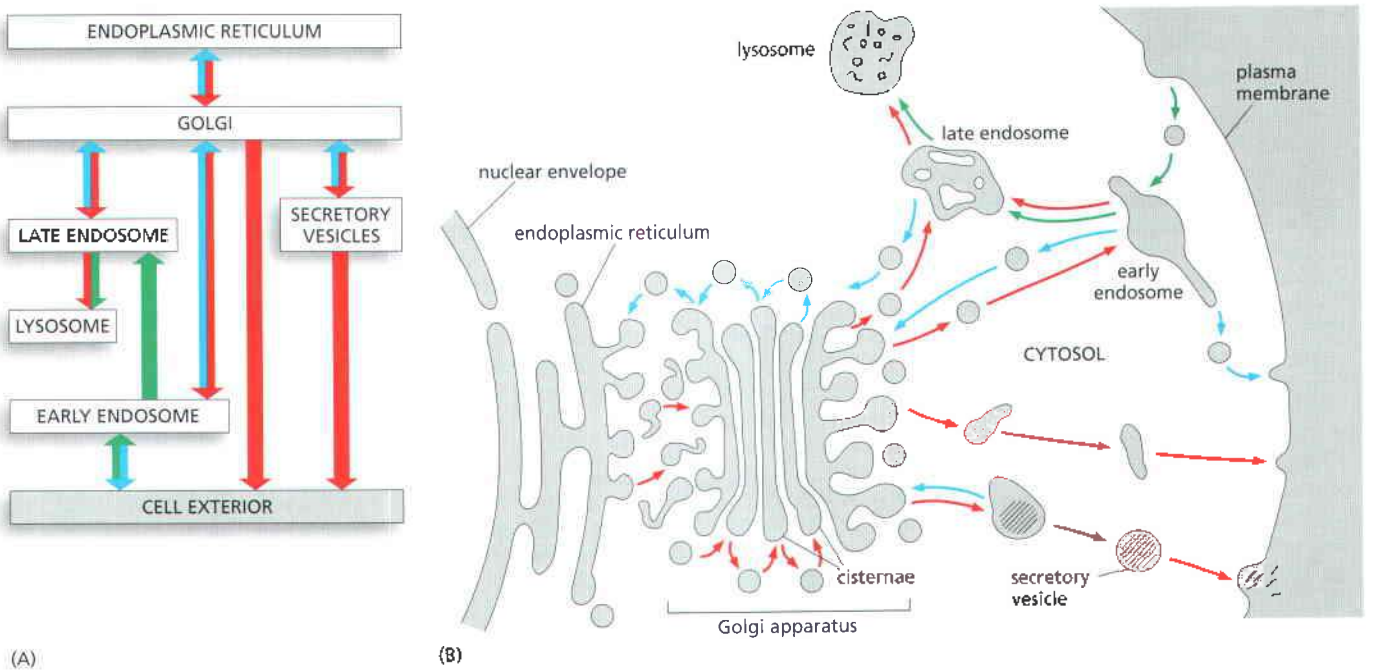


Figure 13–3 **A "road-map" of the biosynthetic–secretory and endocytic pathways.** (A) In this diagram, which was introduced in Chapter 12, the endocytic and biosynthetic–secretory pathways are illustrated with *green* and *red* arrows, respectively. In addition, *blue* arrows denote retrieval pathways for the backflow of selected components. (B) The compartments of the eucaryotic cell involved in vesicular transport. The lumen of each membrane-enclosed compartment is topologically equivalent to the outside of the cell. All compartments shown communicate with one another and the outside of the cell by means of transport vesicles. In the biosynthetic–secretory pathway *(red arrows)*, protein molecules are transported from the ER to the plasma membrane or (via endosomes) to lysosomes. In the endocytic pathway *(green arrows)*, molecules are ingested in vesicles derived from the plasma membrane and delivered to early endosomes and then (via late endosomes) to lysosomes. Many endocytosed molecules are retrieved from early endosomes and returned to the cell surface for reuse; similarly, some molecules are retrieved from the early and late endosomes and returned to the Golgi apparatus, and some are retrieved from the Golgi apparatus and returned to the ER. All of these retrieval pathways are shown with *blue arrows*, as in part (A).

used to extract specific components from a membrane for delivery to another compartment. Finally, we discuss how transport vesicles dock at the appropriate target membrane and fuse with it to deliver their cargo.

There Are Various Types of Coated Vesicles

Most transport vesicles form from specialized, coated regions of membranes. They bud off as **coated vesicles**, which have a distinctive cage of proteins covering their cytosolic surface. Before the vesicles fuse with a target membrane, they discard their coat, as is required for the two cytosolic membrane surfaces to interact directly and fuse.

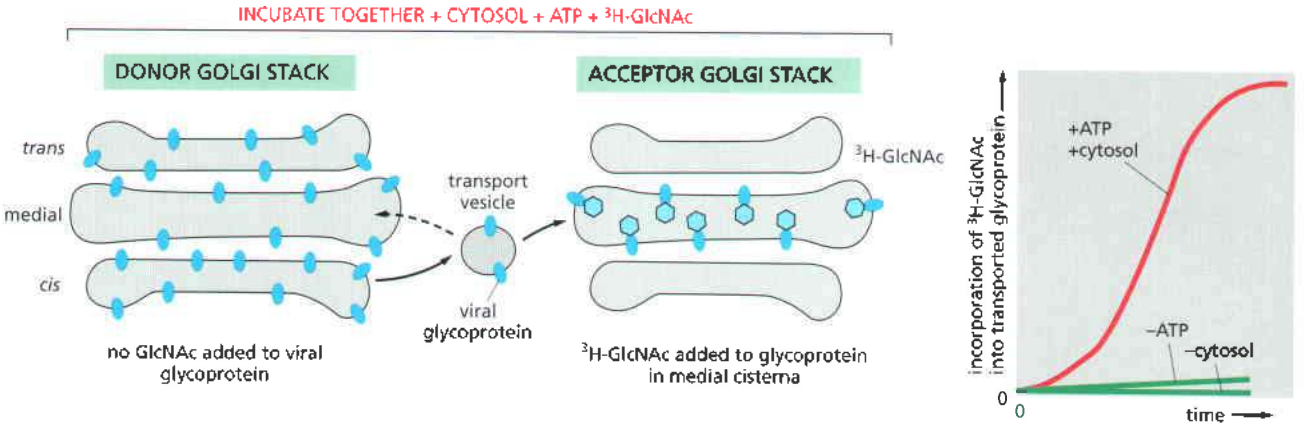
The coat performs two main functions. First, it concentrates specific membrane proteins in a specialized patch, which then gives rise to the vesicle membrane. In this way, it selects the appropriate molecules for transport. Second, the coat molds the forming vesicle. Coat proteins assemble into a curved, basketlike lattice that deforms the membrane patch and thereby shapes the vesicle. This may explain why vesicles with the same type of coat often have a relatively uniform size and shape.

There are three well-characterized types of coated vesicles, distinguished by their coat proteins: *clathrin-coated*, *COPI-coated*, and *COPII-coated* (**Figure 13–4**). Each type is used for different transport steps. Clathrin-coated vesicles, for example, mediate transport from the Golgi apparatus and from the plasma membrane, whereas COPI- and COPII-coated vesicles most commonly mediate transport from the ER and from the Golgi cisternae (**Figure 13–5**). There is, however, much more variety in coated vesicles and their functions than this short list suggests. As we discuss below, there are several types of clathrin-coated vesicles,

CELL-FREE SYSTEMS FOR STUDYING THE COMPONENTS AND MECHANISM OF VESICULAR TRANSPORT

Vesicular transport can be reconstituted in cell-free systems. This was first achieved for the Golgi stack. When Golgi stacks are isolated from cells and incubated with cytosol and with ATP as a source of energy, transport vesicles bud from their rims and appear to transport proteins between cisternae. By following the progressive processing of the oligosaccharides on a glycoprotein as it moves from one Golgi compartment to the next, it is possible to follow the process of vesicular transport.

To follow the transport, two distinct populations of Golgi stacks are incubated together. The "donor" population is isolated from mutant cells that lack the enzyme *N*-acetylglucosamine (GlcNAc) transferase I and that have been infected with a virus; because of the mutation, the major viral glycoprotein fails to be modified with GlcNAc in the Golgi apparatus of the mutant cells. The "acceptor" Golgi stacks are isolated from uninfected wild-type cells and thus contain a good copy of GlcNAc transferase I, but lack the viral glycoprotein. In the mixture of Golgi stacks the viral glycoprotein acquires GlcNAc, indicating that it must have been transported between the Golgi stacks—presumably by vesicles that bud from the *cis* compartment of the donor Golgi and fuse with the medial compartment of the acceptor Golgi or by homotypic fusion between the cisternae of two Golgi stacks. This transport-dependent glycosylation is monitored by measuring the transfer of ^{3}H-GlcNAc from UDP-^{3}H-GlcNAc to the viral glycoprotein. Transport occurs only when ATP and cytosol are added. By fractionating the cytosol, a number of specific cytosolic proteins have been identified that are required for the budding and fusion of transport vesicles.



Similar cell-free systems have been used to study transport from the medial to the *trans* Golgi network, from the *trans* Golgi network to the plasma membrane, from endosomes to lysosomes, from the *trans* Golgi network to late endosomes, and to study homotypic fusion between like compartments—such as endosomes and immature secretory vesicles.

GENETIC APPROACHES FOR STUDYING VESICULAR TRANSPORT

Genetic studies of mutant yeast cells defective for secretion have identified more than 25 genes that are involved in the secretory pathway. Many of the mutant genes encode *temperature-sensitive* proteins. These function normally at 25°C, but when the mutant cells (A–I) are shifted to an elevated temperature, such as 35°C, they fail to transport proteins from the ER to the Golgi apparatus, others from one Golgi cisterna to another, and still others from the Golgi apparatus to the vacuole (the yeast lysosome) or to the plasma membrane.

Once a protein required for secretion has been identified in this way, a phenomenon called *multicopy suppression* can be used to identify genes that encode other proteins that interact with it. At high temperatures, a temperature-sensitive mutant protein often has too low an affinity for its normal interaction partners. If the interacting proteins are produced at much higher concentration than normal, however, sufficient binding occurs to cure the defect. To create an experimental paradigm in which such high concentrations of ligand are present, mutant yeast cells (with a temperature-sensitive mutation in a gene involved in vesicular transport) are transfected with a yeast plasmid vector into which random normal yeast genomic DNA fragments have been cloned. Because these plasmids are maintained in cells at high copy number, any cells that happen to carry plasmids with intact genes will overproduce the normal gene product, allowing rare cells to survive at the high temperature. The relevant DNA fragments, which presumably encode proteins that interact with the original mutant protein, can then be isolated from the surviving cell clones.

The genetic and biochemical approaches complement each other, and many of the proteins involved in vesicular transport have been identified independently by biochemical studies of mammalian cell-free systems and by genetic studies in yeast.

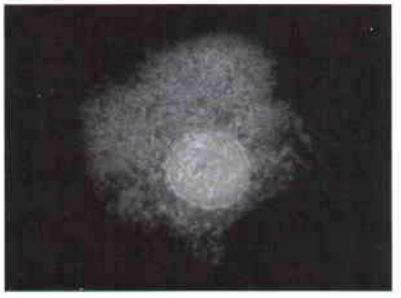
GFP FUSION PROTEINS HAVE REVOLUTIONIZED THE STUDY OF INTRACELLULAR TRANSPORT

One way to follow the whereabouts of a protein in living cells is to construct fusion proteins, in which green fluorescent protein (GFP) is attached by genetic engineering techniques to the protein of interest. When a cDNA encoding such a fusion protein is expressed in a cell, the protein is readily visible in a fluorescent microscope, so that it can be followed in living cells in real time. Fortunately, for most proteins studied the addition of GFP does not perturb the protein's function.

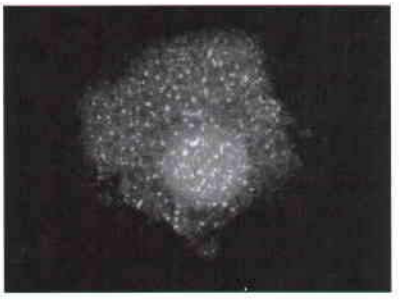
GFP fusion proteins are widely used to study the location and movement of proteins in cells. GFP fused to proteins that shuttle in and out of the nucleus, for example, facilitates studies of nuclear transport and its regulation. GFP fused to mitochondrial or Golgi proteins is used to study the behavior of these organelles. GFP fused to plasma membrane proteins allows measurement of the kinetics of their movement from the ER through the secretory pathway. Dramatic examples of such experiments can be seen as movies on the DVD that accompanies this book.

The study of GFP fusion proteins is often combined with FRAP and FLIP techniques (discussed in Chapter 10), in which the GFP in selected regions of the cell is bleached by strong laser light. The rate of diffusion of unbleached GFP fusion proteins into that area can then be determined to provide measurement of the protein's diffusion or transport in the cell. In this way, for example, it was determined that many Golgi enzymes recycle between the Golgi apparatus and the ER. <GCTG>

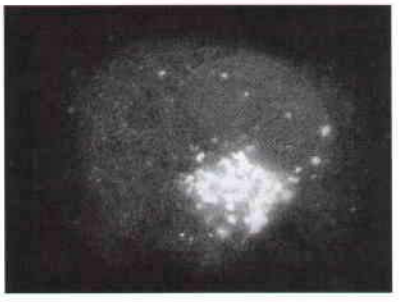
(A–D, courtesy of Jennifer Lippincott-Schwartz Lab.)



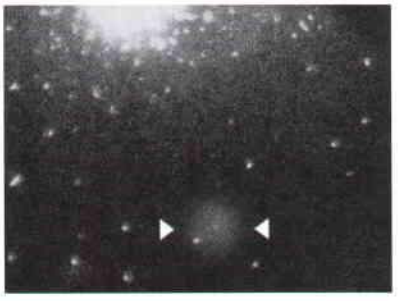
(A) In this experiment, cultured cells express a GFP fusion protein consisting of GFP attached to a viral coat protein—called vesicular stomatitis virus coat protein. The viral protein is an integral membrane protein that normally moves through the secretory pathway from the ER to the cell surface, where the virus would be assembled if cells also expressed the other viral components. The viral protein contains a mutation that allows export from the ER only at a low temperature. Thus, at the high temperature shown, the fusion protein labels the ER.



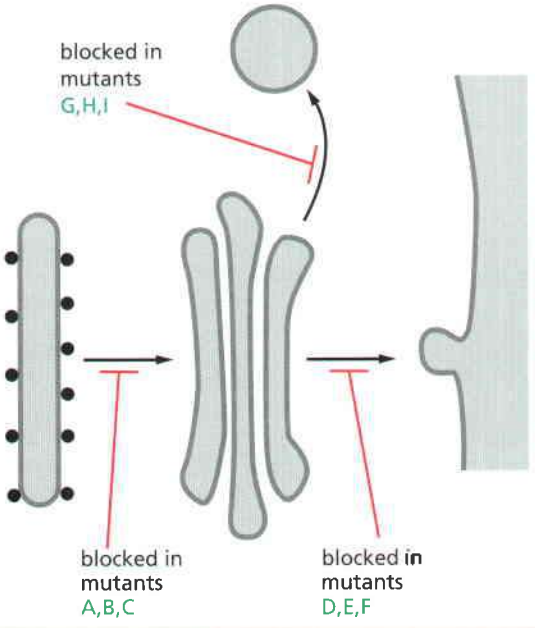
(B) As the temperature is lowered, the GFP fusion protein rapidly accumulates at ER exit sites.



(C) The fusion protein then moves to the Golgi apparatus.



(D) Finally, the fusion protein is delivered to the plasma membrane where the delivered protein diffuses into the plasma membrane (the arrows bracket a fusion event). From such studies the kinetics of each step in the pathway can be determined.



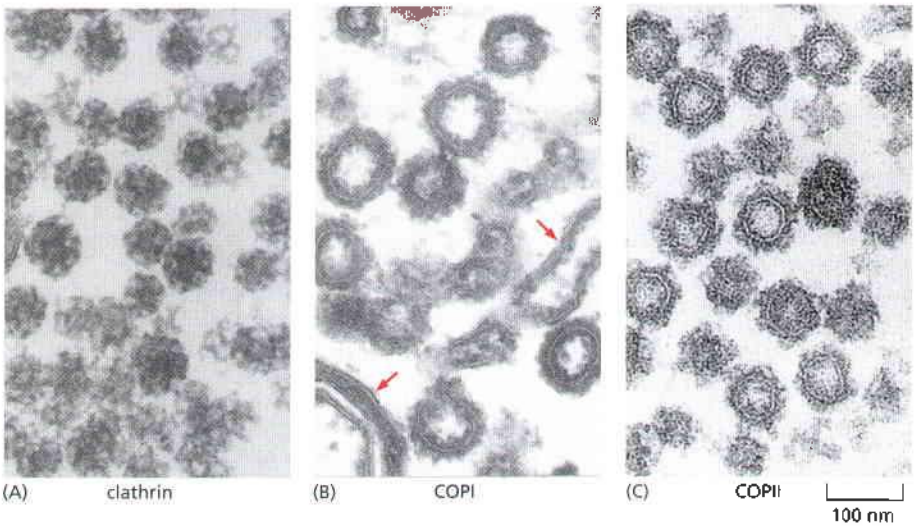


Figure 13–4 Electron micrograph of clathrin-coated, COPI-coated, and COPII-coated vesicles. All are shown in electron micrographs at the same scale. (A) Clathrin-coated vesicles. (B) Golgi cisternae (*arrows*) from a cell-free system in which COPI-coated vesicles bud in the test tube. (C) COPII-coated vesicles. (A and B, from L. Orci, B. Glick and J. Rothman, *Cell* 46:171–184, 1986. With permission from Elsevier; C, courtesy of Charles Barlowe and Lelio Orci.)

each specialized for a different transport step, and the COPI- and COPII-coated vesicles may be similarly diverse.

The Assembly of a Clathrin Coat Drives Vesicle Formation

Clathrin-coated vesicles, the first coated vesicles to be identified, transport material from the plasma membrane and between endosomal and Golgi compartments. **COPI-coated vesicles** and **COPII-coated vesicles** transport material early in the secretory pathway: COPII-coated vesicles bud from the ER, and COPI-coated vesicles bud from Golgi compartments (see Figure 13–5). We discuss clathrin-coated vesicles first, as they provide a good example of how vesicles form.

The major protein component of clathrin-coated vesicles is **clathrin** itself. Each clathrin subunit consists of three large and three small polypeptide chains that together form a three-legged structure called a *triskelion*. Clathrin triskelions assemble into a basketlike convex framework of hexagons and pentagons to form coated pits on the cytosolic surface of membranes (**Figure 13–6**). Under appropriate conditions, isolated triskelions spontaneously self-assemble into typical polyhedral cages in a test tube, even in the absence of the membrane vesicles that these baskets normally enclose (**Figure 13–7**). Thus, the clathrin triskelions determine the geometry of the clathrin cage.

Adaptor proteins, another major coat component in clathrin-coated vesicles, form a discrete second layer of the coat, positioned between the clathrin cage and the membrane. They bind the clathrin coat to the membrane and trap

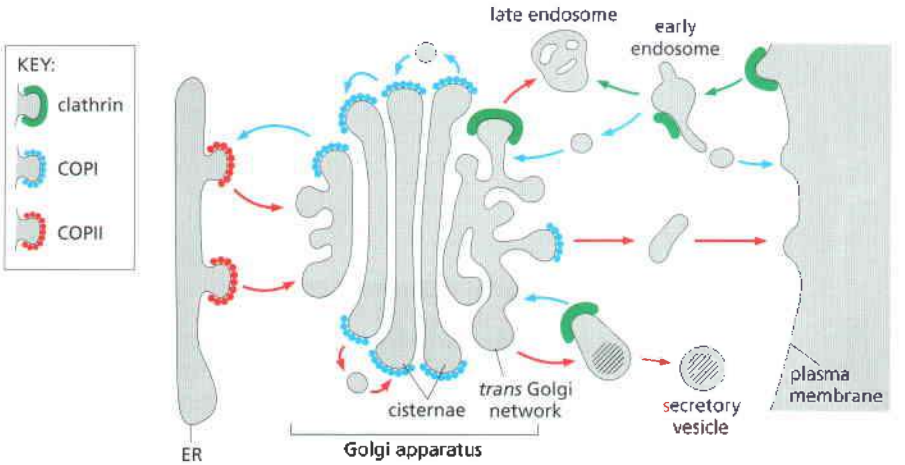


Figure 13–5 Use of different coats in vesicular traffic. Different coat proteins select different cargo and shape the transport vesicles that mediate the various steps in the biosynthetic– secretory and endocytic pathways. When the same coats function in different places in the cell, they usually incorporate different coat protein subunits that modify their properties (not shown). Many differentiated cells have additional pathways beside those shown here, including a sorting pathway from the *trans* Golgi network to the apical surface of epithelial cells and a specialized recycling pathway for proteins of synaptic vesicles in the synapses of neurons.

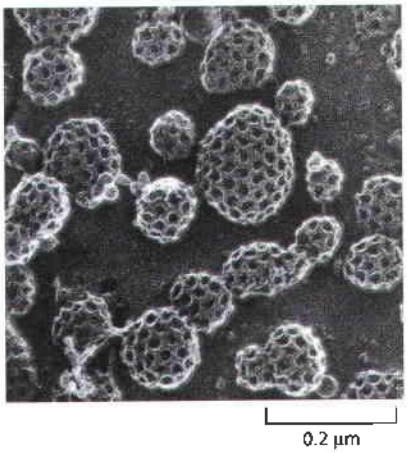


Figure 13–6 Clathrin-coated pits and vesicles. This rapid-freeze, deep-etch electron micrograph shows numerous clathrin-coated pits and vesicles on the inner surface of the plasma membrane of cultured fibroblasts. The cells were rapidly frozen in liquid helium, fractured, and deep-etched to expose the cytoplasmic surface of the plasma membrane. (From J. Heuser, *J. Cell Biol.* 84:560–583, 1980. With permission from The Rockefeller University Press.)

various transmembrane proteins, including transmembrane receptors that capture soluble cargo molecules inside the vesicle—so-called *cargo receptors*. In this way, a selected set of transmembrane proteins, together with the soluble proteins that interact with them, are packaged into each newly formed clathrin-coated transport vesicle (Figure 13–8).

There are several types of adaptor proteins. The best characterized have four different protein subunits; others are single-chain proteins. Each type of adaptor protein is specific for a different set of cargo receptors and its use leads to the formation of distinct clathrin-coated vesicles. Clathrin-coated vesicles budding from different membranes use different adaptor proteins and thus package different receptors and cargo molecules.

The sequential assembly of adaptor complexes and the clathrin coat on the cytosolic surface of the membrane generate forces that result in the formation of a clathrin-coated vesicle. Lateral interactions between adaptor complexes and between clathrin molecules aid in forming the vesicle.

Not All Coats Form Basketlike Structures

Not all coats are as regular and universal as the examples of clathrin or COP coats suggest. Some coats may be better described as specialized protein assemblies that form patches dedicated to specific cargo proteins. An example is a coat called **retromer**, which assembles on endosomes and forms vesicles that return *acid hydrolase receptors*, such as the *mannose-6-phosphate receptor*, to the Golgi apparatus (Figure 13–9). We discuss later the important role of these receptors in delivering enzymes into new lysosomes.

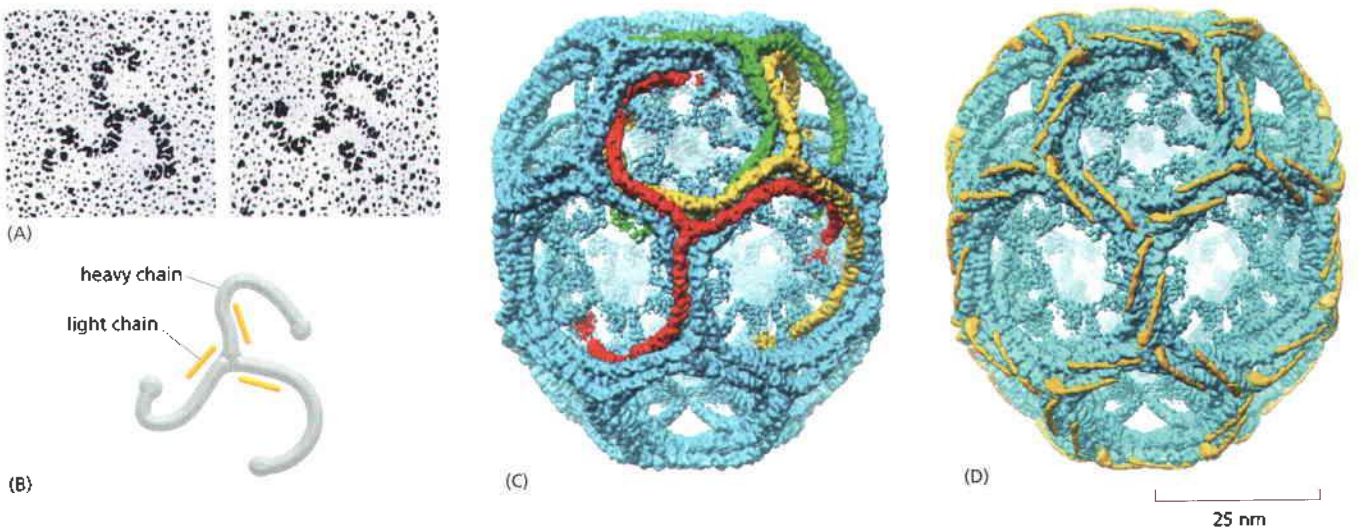


Figure 13–7 The structure of a clathrin coat. <TATT> (A) Electron micrographs of clathrin triskelions shadowed with platinum. Each triskelion is composed of three clathrin heavy chains and three clathrin light chains as shown in (B). (C and D) A cryoelectron micrograph taken of a clathrin coat composed of 36 triskelions organized in a network of 12 pentagons and 6 hexagons, with heavy chains (C) and light chains (D) highlighted. The interwoven legs of the clathrin triskelions form an outer shell into which the N-terminal domains of the triskelions protrude to form an inner layer visible through the openings. It is this inner layer that contacts the adaptor proteins shown in Figure 13–8. Although the coat shown is too small to enclose a membrane vesicle, the clathrin coats on vesicles are constructed in a similar way, from 12 pentagons and a larger number of hexagons, resembling the architecture of a soccer ball. (A, from E. Ungewickell and D. Branton, *Nature* 289:420–422, 1981; C and D, from A. Fotin et al., *Nature* 432:573–579, 2004. All with permission from Macmillan Publishers Ltd.)

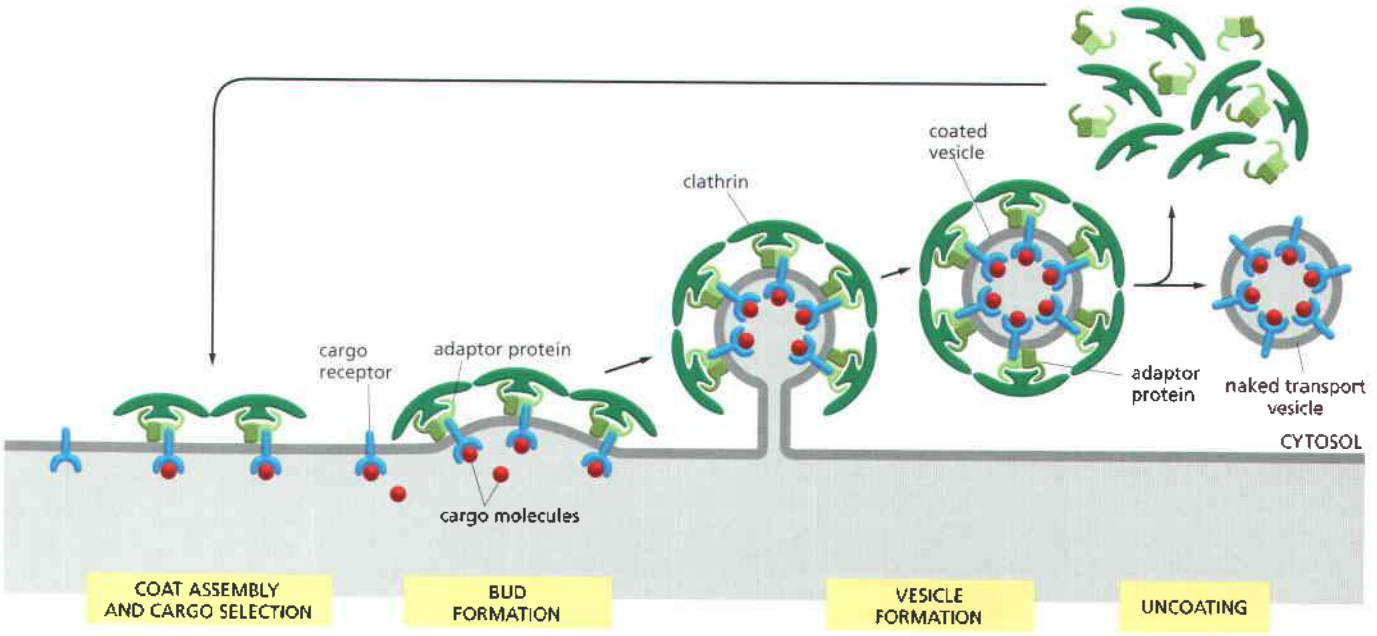


Figure 13–8 **The assembly and disassembly of a clathrin coat.** The assembly of the coat introduces curvature into the membrane, which leads in turn to the formation of uniformly sized coated buds. The adaptor proteins bind clathrin triskelions and membrane-bound cargo receptors, thereby mediating the selective recruitment of both membrane and cargo molecules into the vesicle. The clathrin coat is rapidly lost shortly after the vesicle forms.

Retromer is a multiprotein complex that assembles into a coat on endosomal membranes only when:

1. it can bind to the cytoplasmic tails of the cargo receptors,
2. it can interact directly with a curved phospholipid bilayer, and
3. it can bind to a specific phosphorylated phosphatidylinositol lipid (a *phosphoinositide*), which acts as an endosomal marker, as we discuss next.

Because these three requirements must be met simultaneously, retromer is thought to act as a *coincidence detector* and only assemble at the right time and place. Upon binding as a dimer, it stabilizes the membrane curvature, which makes the binding of additional retromers in its proximity more likely. The cooperative assembly of retromer then leads to the formation and budding of a transport vesicle, which delivers its cargo to the Golgi apparatus.

The adaptor proteins found in clathrin coats also bind to phosphoinositides, which not only have a major role in directing when and where coats assemble in the cell, but are used much more widely as molecular markers of compartment identity. This helps to control membrane trafficking events, as we discuss next.

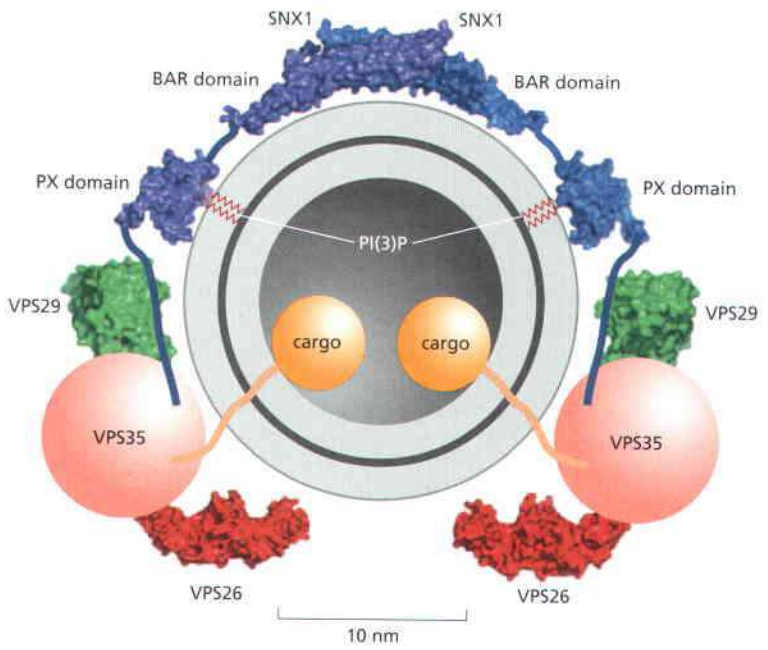


Figure 13–9 **A model for retromer assembly on endosomal membranes.** The four retromer subunits, SNX1, VPS29, VPS35, and VPS26, form coated domains on endosome membranes that capture cargo molecules, including transmembrane proteins such as acid hydrolase receptors, into vesicles that return to the *trans* Golgi network. VPS35 binds to the cytoplasmic tails of the transmembrane cargo proteins. SNX1 protein contains different protein modules: a *PX domain* that binds to the phosphorylated phosphatidylinositol PI(3)P, and a *BAR domain* that mediates dimerization and attachment to highly curved membranes. Both PX and BAR domains are protein modules that are found in many proteins, where they carry out similar functions. With the exception of the PI(3)P, which is enlarged for visibility, the membrane and other components are drawn approximately to scale. (Adapted from J.S. Bonifacino and R. Rojas, *Nat. Rev. Mol. Cell Biol.* 7:568–579, 2006. With permission from Macmillan Publishers Ltd.)

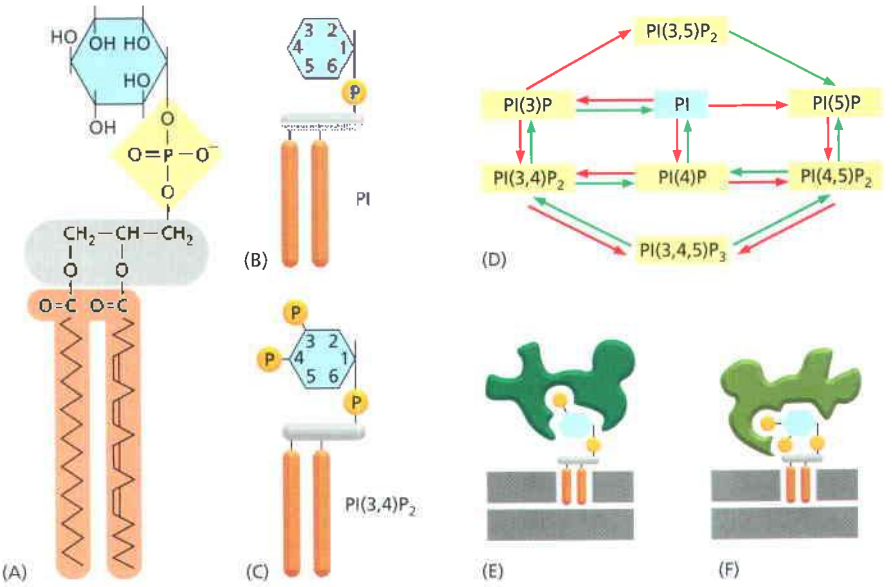


Figure 13–10 Phosphatidylinositol (PI) and phosphoinositides (PIPs). (A, B) The structure of PI shows the free hydroxyl groups in the inositol sugar that can in principle be modified. (C) Phosphorylation of one, two, or three of the hydroxyl groups on PI by PI and PIP kinases produces a variety of PIP species. They are named according to the ring position (in parentheses) and the number of phosphate groups (subscript) added to PI. $PI(3,4)P_2$ is shown. (D) Animal cells have several PI and PIP kinases and a similar number of PIP phosphatases, which are localized to different organelles, where they are regulated to catalyze the production of particular PIPs. The *red* and *green* arrows show the kinase and phosphatase reactions, respectively. (E, F) Phosphoinositide head groups are recognized by protein domains that discriminate between the different forms. In this way, selected groups of proteins containing such domains are recruited to regions of membrane in which these phosphoinositides are present. PI(3)P and $PI(4,5)P_2$ are shown. (D, modified from M.A. de Matteis and A. Godi, *Nat. Cell Biol.* 6:487–492, 2004. With permission from Macmillan Publishers Ltd.)

Phosphoinositides Mark Organelles and Membrane Domains

Although inositol phospholipids typically comprise less than 10% of the total phospholipids in a membrane, they have important regulatory functions. They can undergo rapid cycles of phosphorylation and dephosphorylation at the 3′, 4′, and 5′ positions of their inositol sugar head groups to produce various types of **phosphoinositides (PIPs)**. The interconversion of phosphatidylinositol (PI) and PIPs is highly compartmentalized: different organelles in the endocytic and biosynthetic–secretory pathways have distinct sets of PI and PIP kinases and PIP phosphatases (**Figure 13–10**). The distribution, regulation, and local balance of these enzymes determine the steady-state distribution of each PIP species. As a consequence, the distribution of PIPs varies from organelle to organelle, and often within a continuous membrane from one region to another, thereby defining specialized membrane domains.

Many proteins involved at different steps in vesicular transport contain domains that bind with high specificity to the head groups of particular PIPs, distinguishing one phosphorylated form from another. Local control of the PI and PIP kinases and PIP phosphatases can therefore be used to rapidly control the binding of proteins to a membrane or membrane domain. The production of a particular type of PIP recruits proteins containing matching PIP-binding domains. The PIP-binding proteins then help regulate vesicle formation and other steps in membrane transport (**Figure 13–11**). The same strategy is widely used to recruit specific intracellular signaling proteins to the plasma membrane in response to extracellular signals (discussed in Chapter 15).

Cytoplasmic Proteins Regulate the Pinching-Off and Uncoating of Coated Vesicles

As a clathrin-coated bud grows, soluble cytoplasmic proteins, including **dynamin**, assemble as a ring around the neck of each bud (**Figure 13–12**). Dynamin contains a $PI(4,5)P_2$-binding domain, which tethers the protein to the membrane, and a GTPase domain, which regulates the rate at which vesicles pinch off from the membrane. The pinching-off process brings the two noncytosolic leaflets of the membrane into close proximity and fuses them, sealing off the forming vesicle. To perform this task, dynamin recruits other proteins to the neck of the budding vesicle, which, together with dynamin, help to bend the patch of membrane—by directly distorting the bilayer structure, by changing its lipid composition through the recruitment of lipid-modifying enzymes, or by both mechanisms.

Once released from the membrane, the vesicle rapidly loses its clathrin coat. A PIP phosphatase that is co-packaged into clathrin-coated vesicles depletes

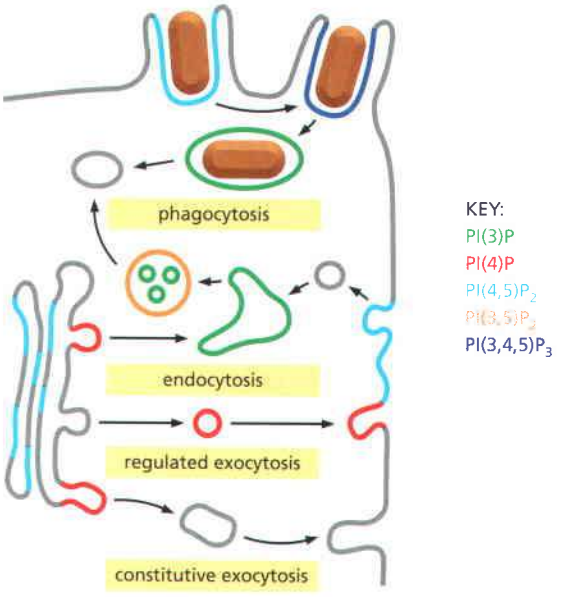
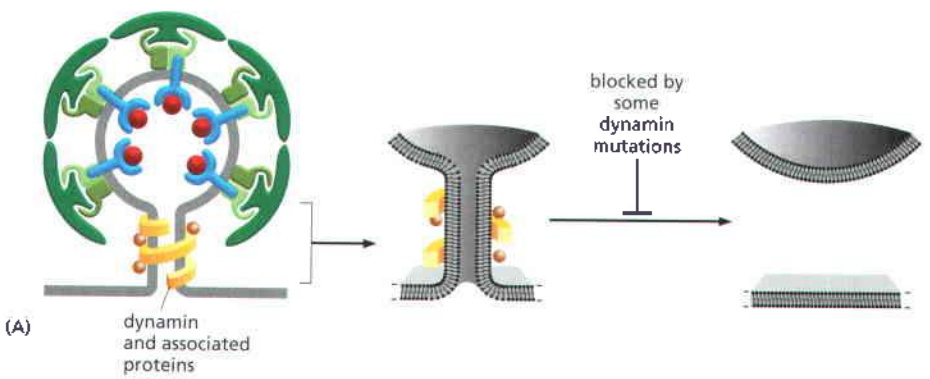


Figure 13–11 The intracellular location of phosphoinositides. Different types of PIPs are located in different membranes and membrane domains, where they are often associated with specific vesicular transport events. The membrane of secretory vesicles, for example, contains PI(4)P. When the vesicles fuse with the plasma membrane, a PI 5-kinase that is localized there converts the PI(4)P into PI(4,5)P_2. The PI(4,5)P_2, in turn, helps recruit adaptor proteins, which initiate the formation of a clathrin-coated pit, as the first step in clathrin-mediated endocyosis. Once the clathrin-coated vesicle buds off from the plasma membrane, a PI(5)P phosphatase hydrolyzes PI(4,5)P_2, which weakens the binding of the adaptor proteins, promoting vesicle uncoating. We discuss phagocytosis and the distinction between regulated and constitutive exocytosis later in the chapter. (Modified from M.A. de Matteis and A. Godi, *Nat. Cell Biol.* 6:487–492, 2004. With permission from Macmillan Publishers Ltd.)

PI(4,5)P_2 from the membrane, which weakens the binding of the adaptor proteins. In addition, an Hsp70 chaperone protein functions as an uncoating ATPase, using the energy of ATP hydrolysis to peel off the clathrin coat. *Auxillin*, another vesicle protein, is believed to activate the ATPase. Because the coated bud persists much longer than the coat on the vesicle, additional control mechanisms must somehow prevent the clathrin from being removed before it has formed a complete vesicle (discussed below).

Monomeric GTPases Control Coat Assembly

To balance the vesicular traffic to and from a compartment, coat proteins must assemble only when and where they are needed. While local production of PIPs plays a major part in regulating the assembly of clathrin coats on the plasma membrane and Golgi apparatus, cells have additional ways of regulating coat formation. *Coat-recruitment GTPases*, for example, control the assembly of clathrin coats on endosomes and the COPI and COPII coats on Golgi and ER membranes.



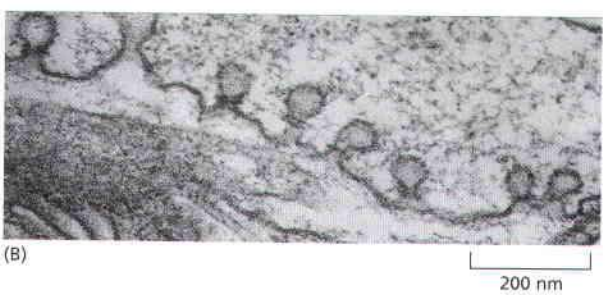


Figure 13–12 The role of dynamin in pinching off clathrin-coated vesicles. (A) The dynamin assembles into a ring around the neck of the forming bud. The dynamin ring is thought to recruit other proteins to the vesicle neck, which, together with dynamin, destabilize the interacting lipid bilayers so that the noncytoplasmic leaflets flow together. The newly formed vesicle then pinches off from the membrane. Specific mutations in dynamin can either enhance or block the pinching-off process. (B) Dynamin was discovered as the protein defective in the *shibire* mutant of *Drosophila*. These mutant flies become paralyzed because clathrin-mediated endocytosis stops, and the synaptic vesicle membrane fails to recycle, blocking neurotransmitter release. Deeply invaginated clathrin-coated pits form in the fly's nerve cells, with a ring of mutant dynamin assembled around the neck, as shown in this thin-section electron micrograph. The pinching-off process fails because membrane fusion does not take place. (B, from J.H. Koenig and K. Ikeda, *J. Neurosci.* 9:3844–3860, 1989. With permission from the Society of Neuroscience.)

Many steps in vesicular transport depend on a variety of GTP-binding proteins that control both the spatial and temporal aspects of membrane exchange. As discussed in Chapter 3, GTP-binding proteins regulate most processes in eucaryotic cells. They act as molecular switches, which flip between an active state with GTP bound and an inactive state with GDP bound. Two classes of proteins regulate the flipping: *guanine-nucleotide-exchange factors (GEFs)* activate the proteins by catalyzing the exchange of GDP for GTP, and *GTPase-activating proteins (GAPs)* inactivate the proteins by triggering the hydrolysis of the bound GTP to GDP (see Figure 3–71) Although both monomeric GTP-binding proteins (monomeric GTPases) and trimeric GTP-binding proteins (G proteins) have important roles in vesicular transport, the roles of the monomeric GTPases are better understood, and we focus on them here.

Coat-recruitment GTPases are members of a family of monomeric GTPases. They include the **Arf proteins**, which are responsible for both COPI coat assembly and clathrin coat assembly at Golgi membranes, and the **Sar1 protein**, which is responsible for COPII coat assembly at the ER membrane. Coat-recruitment GTPases are usually found in high concentration in the cytosol in an inactive, GDP-bound state. When a COPII-coated vesicle is to bud from the ER membrane, a specific Sar1-GEF embedded in the ER membrane binds to cytosolic Sar1, causing the Sar1 to release its GDP and bind GTP in its place. (Recall that GTP is present in much higher concentration in the cytosol than GDP and therefore will spontaneously bind after GDP is released.) In its GTP-bound state, the Sar1 protein exposes an amphiphilic helix, which inserts into the cytoplasmic leaflet of the lipid bilayer of the ER membrane. The tightly bound Sar1 now recruits coat protein subunits to the ER membrane to initiate budding (**Figure 13–13**). Other GEFs and coat-recruitment GTPases operate in a similar way on other membranes.

Some coat protein subunits also interact, albeit more weakly, with the head groups of certain lipid molecules, in particular phosphatidic acid and phosphoinositides, as well as with the cytoplasmic tail of some of the transmembrane

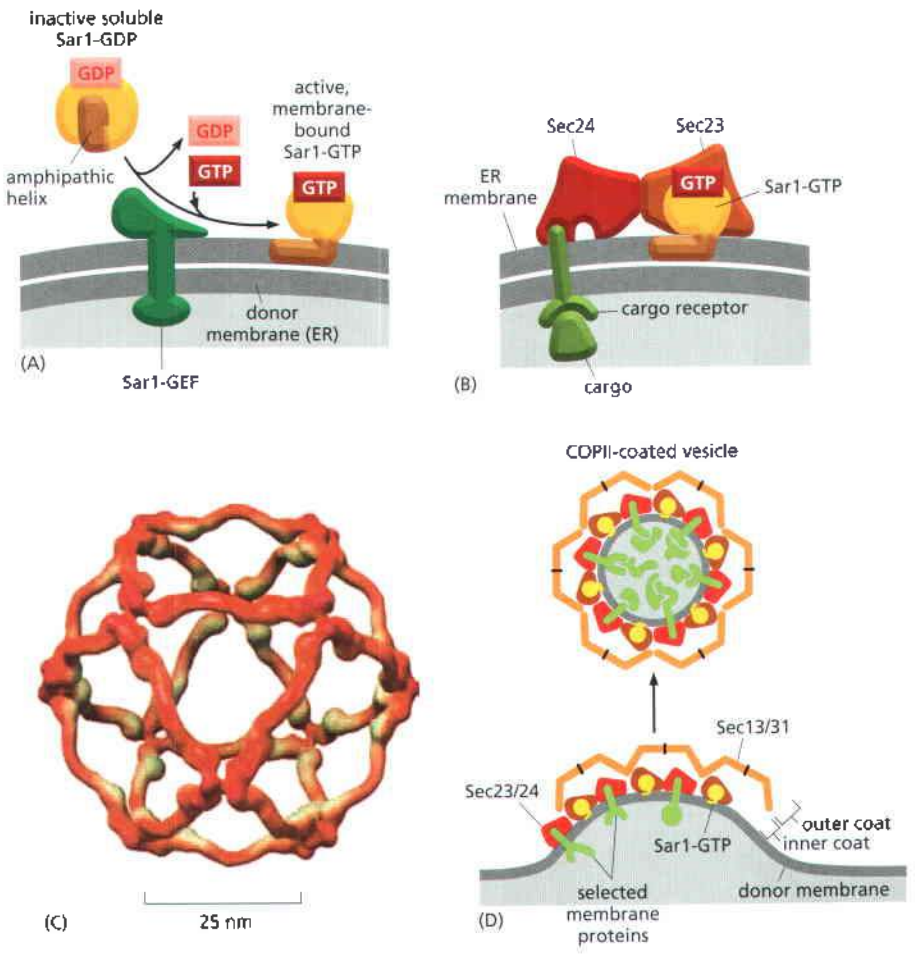


Figure 13–13 Formation of a COPII-coated vesicle. (A) The Sar1 protein is a coat-recruitment GTPase. Inactive, soluble Sar1-GDP binds to a Sar1-GEF in the ER membrane, causing the Sar1 to release its GDP and bind GTP. The Sar1-GEF (also called Sec12) was discovered as one of the temperature-sensitive mutants that block ER to Golgi transport (see Panel 13–1). A GTP-triggered conformational change in Sar1 exposes an amphipathic helix, which inserts into the cytoplasmic leaflet of the ER membrane, initiating a process of membrane curvature. (B) GTP-bound Sar1 binds to a complex of two COPII coat proteins, called Sec23/24. The crystal structure of the Sec23/24 proteins together with Sar1 predicts how the amphiphilic helix of Sar1 anchors the complex to the membrane. Sec24 has several different binding sites for cytoplasmic tails of cargo receptors. The entire surface of the complex that attaches to the membrane is gently curved to match the diameter of COPII-coated vesicles. (C) A complex of two additional COPII coat proteins, called Sec13/31, forms the outer shell of the coat. Like clathrin, Sec13/31 alone can assemble into symmetrical cages with appropriate dimensions to enclose a COPII-coated vesicle. (D) Membrane-bound, active Sar1-GTP recruits COPII subunits to the membrane. This causes the membrane to form a bud, which includes selected transmembrane proteins. A subsequent membrane-fusion event pinches off the coated vesicle.

Other coated vesicles are thought to form in a similar way. The coat-recruitment GTPase Arf also contains a regulated amphiphilic helix but, in contrast to Sar1, the helix also has a covalently attached fatty acid chain that contributes to its hydrophobicity. As in Sar1, the regulated amphiphilic helix is retracted in the GDP-bound state and exposed in the GTP-bound state. As we discuss later, *Rab GTPases* regulate their membrane attachment in a similar manner (see Figure 13–14). (C, modified from S.M. Stagg et al., *Nature* 439:234–238, 2006. With permission from Macmillan Publishers Ltd.)

proteins they recruit into the bud. Together, these protein–lipid and protein–protein and interactions tightly bind the coat to the membrane, causing the membrane to deform into a bud, which then pinches off as a coated vesicle.

The coat-recruitment GTPases also have a role in coat disassembly. The hydrolysis of bound GTP to GDP causes the GTPase to change its conformation so that its hydrophobic tail pops out of the membrane, causing the vesicle's coat to disassemble. Although it is not known what triggers the GTP hydrolysis process, it has been proposed that the GTPases work like timers, which hydrolyze GTP at a slow but predictable rate. COPII coats, for example, accelerate GTP hydrolysis by Sar1, thereby triggering coat disassembly at a certain time after coat assembly has begun. Thus, a fully formed vesicle will be produced only when bud formation occurs faster than the timed disassembly process; otherwise, disassembly will be triggered before a vesicle pinches off, and the process will have to start again at a more appropriate time and place.

Not All Transport Vesicles Are Spherical

Although vesicle budding at various locations in the cell has many similarities, each cell membrane poses its own special challenges. The plasma membrane, for example, is comparatively flat and stiff, owing to its cholesterol-rich lipid composition and underlying cortical cytoskeleton. Thus, clathrin coats have to produce considerable force to introduce curvature, especially at the neck of the bud where dynamin and its associated proteins facilitate the sharp bends required for the pinching-off process. In contrast, vesicle budding from many intracellular membranes occurs preferentially at regions where the membranes are already curved, such as the rims of the Golgi cisternae or ends of membrane tubules. In these places, the primary function of the coats is to capture the appropriate cargo proteins rather than to deform the membrane.

Transport vesicles occur in various sizes and shapes. When living cells are genetically engineered to express fluorescent membrane components, the endosomes and *trans* Golgi network are seen in a fluorescence microscope to continually send out long tubules. Coat proteins assemble onto the tubules and help recruit specific cargo. The tubules then either regress or pinch off with the help of dynamin-like proteins to form transport vesicles. Depending on the relative efficiencies of tubule formation and the pinching-off process, vesicles of different sizes and shapes are produced. Thus, vesicular transport does not necessarily occur only through uniformly sized spherical vesicles, but can involve larger portions of a donor organelle.

Tubules have a much higher surface-to-volume ratio than the organelles from which they form. They are therefore relatively enriched in membrane proteins compared with soluble cargo proteins. As we discuss later, this property of tubules is used for sorting proteins in endosomes.

Rab Proteins Guide Vesicle Targeting

To ensure an orderly flow of vesicular traffic, transport vesicles must be highly selective in recognizing the correct target membrane with which to fuse. Because of the diversity and crowding of membrane systems in the cytoplasm, a vesicle is likely to encounter many potential target membranes before it finds the correct one. Specificity in targeting is ensured because all transport vesicles display surface markers that identify them according to their origin and type of cargo, and target membranes display complementary receptors that recognize the appropriate markers. This crucial process depends on two types of proteins: *Rab proteins* direct the vesicle to specific spots on the correct target membrane, and then *SNARE* proteins mediate the fusion of the lipid bilayers.

Rab proteins play a central part in the specificity of vesicular transport. Like the coat-recruitment GTPases discussed earlier (see Figure 13–13), they also are monomeric GTPases. With over 60 known members, they are the largest subfamily of such GTPases. Each Rab protein is associated with one or more membrane-

Table 13–1 Subcellular Locations of Some Rab Proteins

PROTEIN	ORGANELLE
Rab1	ER and Golgi complex
Rab2	*cis* Golgi network
Rab3A	synaptic vesicles, secretory granules
Rab4/Rab11	recycling endosomes
Rab5A	plasma membrane, clathrin-coated vesicles, early endosomes
Rab5C	early endosomes
Rab6	medial and *trans* Golgi cisternae
Rab7	late endosomes
Rab8	early endosomes
Rab9	late endosomes, *trans* Golgi network

enclosed organelles of the biosynthetic–secretory or endocytic pathways, and each of these organelles has at least one Rab protein on its cytosolic surface (**Table 13–1**). Their highly selective distribution on these membrane systems makes Rab proteins ideal molecular markers for identifying each membrane type and guiding vesicular traffic between them. Rab proteins can function on transport vesicles, on target membranes, or both.

Like the coat-recruitment GTPases, Rab proteins cycle between a membrane and the cytosol and regulate the reversible assembly of protein complexes on the membrane. In their GDP-bound state they are inactive and bound to another protein (*Rab-GDP dissociation inhibitor* or *GDI*) that keeps them soluble in the cytosol, while in their GTP-bound state they are active and tightly associated with the membrane of an organelle or transport vesicle. Membrane bound Rab-GEFs activate Rab proteins on both transport vesicle and target membranes, as activate Rab molecules are usually required on both sides. Once in its GTP-bound state and membrane-bound through a hydrophobic lipid anchor, Rab proteins bind to other proteins, called **Rab effectors**, which facilitate vesicle transport, membrane tethering, and fusion (**Figure 13–14**). The GTP hydrolysis sets the concentration of active Rab and, consequently, the concentration of its effectors on the membrane.

In contrast to the highly conserved structure of Rab proteins, the structures of Rab effectors vary greatly. Some Rab effectors, for example, are *motor proteins* that propel vesicles along actin filaments or microtubules to their target membrane. Others are *tethering proteins*, some of which have long threadlike domains that serve as "fishing lines" that can extend to link two membranes more than 200 nm apart; other tethering proteins are large protein complexes that link two membranes that are closer together. Rab effectors can also interact with SNAREs, thus coupling membrane tethering to fusion.

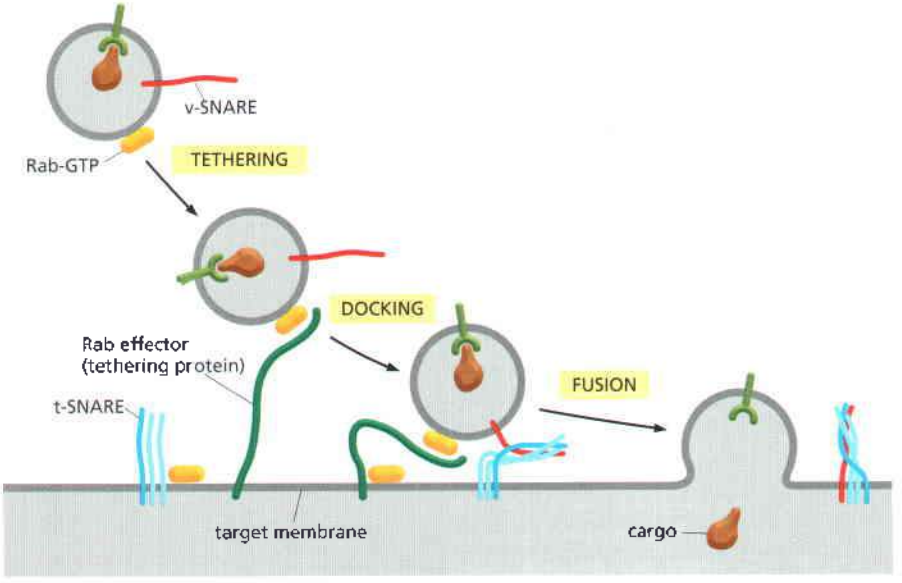


Figure 13–14 Tethering of a vesicle to a target membrane. Rab effector proteins interact via active Rab proteins (Rab-GTPs, *yellow*) located on the target membrane, vesicle membrane, or both, to establish the first connection between the two membranes that are going to fuse. In the example shown here, the Rab effector is a filamentous tethering protein (*green*). Next, SNARE proteins on the two membranes (*red* and *blue*) pair to dock the vesicle to the target membrane and catalyze the fusion of the two apposed lipid bilayers.

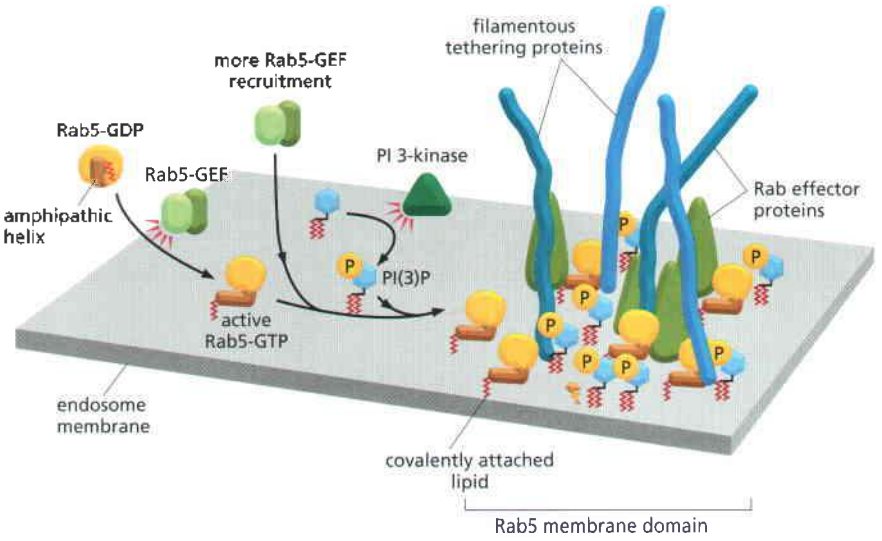


Figure 13–15 The formation of a Rab5 domain on the endosome membrane. A Rab5-specific GEF in the endosome membrane binds a Rab5 protein and induces it to exchange GDP for GTP. GTP binding alters the conformation of the Rab protein, exposing an amphiphilic helix and a covalently attached lipid group, which together anchor the Rab5-GTP to the membrane. Active Rab5 activates PI 3-kinase, which converts PI into PI(3)P. PI(3)P and active Rab5 together bind a variety of Rab effector proteins that contain PI(3)P-binding sites, including filamentous tethering proteins that catch incoming clathrin-coated vesicles from the plasma membrane. Active Rab5 also recruits more Rab5-GEF, further enhancing the assembly of the Rab5 domain on the membrane.

Controlled cycles of GTP hydrolysis and GDP–GTP exchange are thought to regulate the size and activity of such Rab domains dynamically. Unlike SNAREs, which are integral membrane proteins, the GDP/GTP cycle coupled to the membrane/cytosol translocation cycle endows the Rab machinery with the ability to undergo assembly and disassembly on the membrane. For example, during transport from early to late endosomes, Rab5 can be moved and replaced by Rab7, thus committing cargo to degradation. (Adapted from M. Zerial and H. McBride, *Nat. Rev. Mol. Cell Biol.* 2:107–117, 2001. With permission from Macmillan Publishers Ltd.)

The same Rab proteins can bind to multiple effectors. The assembly of Rab proteins and their effectors on a membrane is cooperative and results in the formation of large, specialized membrane patches. Rab5, for example, assembles on endosomal membranes and mediates the capture of clathrin-coated vesicles arriving from the plasma membrane. It recruits tethering proteins to catch the incoming vesicles. Rab5-GEF initially recruits Rab5 to the endosome and converts it to its active GTP-bound form, which becomes anchored to the membrane (**Figure 13–15**). Active Rab5 recruits more Rab5-GEF to the endosome, thereby stimulating the recruitment of more Rab5 to the same site. In addition, active Rab5 activates a PI 3-kinase, which locally converts PI to PI(3)P, which in turn binds some of the Rab effectors. This type of positive feedback greatly enhances the assembly process and helps to establish functionally distinct membrane domains within a continuous membrane.

The endosomal membrane provides a striking example of how different Rab proteins and their effectors help to create multiple specialized membrane domains, each fulfilling a particular set of functions. Thus, while the Rab5 domain receives incoming vesicles from the plasma membrane, distinct Rab11 and Rab4 domains in the same membrane are thought to organize the budding of recycling vesicles that return proteins from the endosome to the plasma membrane. Once assembled, these domains coexist in the same membrane over extended periods.

SNAREs Mediate Membrane Fusion

Once a transport vesicle has been tethered to its target membrane, it unloads its cargo by membrane fusion. Membrane fusion requires bringing the lipid bilayers of two membranes to within 1.5 nm of each other so that they can join. When the membranes are in such close apposition, lipids can flow from one bilayer to the other. For this close approach, water must be displaced from the hydrophilic surface of the membrane—a process that is highly energetically unfavorable. It seems likely that specialized *fusion proteins* that overcome this energy barrier catalyze all membrane fusions in cells. We have already discussed the role of dynamin in a related task during the pinching-off of clathrin-coated vesicles (see Figure 13–12).

The **SNARE proteins** (also called **SNAREs**, for short) catalyze the membrane fusion reactions in vesicular transport. They also provide an additional layer of specificity in the transport process by helping to ensure that only correctly targeted vesicles fuse. There are at least 35 different SNAREs in an animal cell, each associated with a particular organelle in the biosynthetic–secretory or endocytic pathway. These transmembrane proteins exist as complementary sets—with **v-SNAREs** usually found on vesicle membranes and **t-SNAREs** usually found on target membranes (see Figure 13–14). A v-SNARE is a single polypeptide chain,

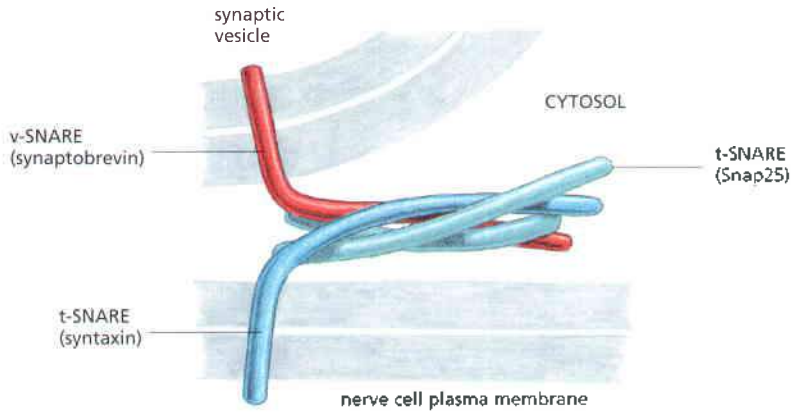


Figure 13–16 The structure of a trans-SNARE complex. The SNAREs responsible for docking synaptic vesicles at the plasma membrane of nerve terminals consist of three proteins. The v-SNARE *synaptobrevin* and the t-SNARE *syntaxin* are both transmembrane proteins, and each contributes one α helix to the complex. The t-SNARE *Snap25* is a peripheral membrane protein that contributes two α helices to the four-helix bundle. They are connected by a loop (omitted from the figure) that lies parallel to the membrane and has fatty acyl chains attached to anchor it there. In many other t-SNARES, each α helix is a separate protein anchored in the membrane by a transmembrane helix. Trans-SNARE complexes always consist of four tightly intertwined α helices, three contributed by a t-SNARE and one by a v-SNARE. The t-SNAREs are composed of multiple chains, one of which is always a transmembrane protein and contributes one helix, and one or two additional light chains that may or may not be transmembrane proteins and that contribute the remaining two helices to the four-helix bundle of the trans-SNARE complex. The crystal structure of a stable complex of the four intertwining α helices contributed by these proteins is modeled here in the context of the whole proteins. The α helices are shown as rods for simplicity. (Adapted from R.B. Sutton et al., *Nature* 395:347–353, 1998. With permission from Macmillan Publishers Ltd.)

whereas a t-SNARE is composed of two or three proteins (**Figure 13–16**). The v-SNAREs and t-SNAREs have characteristic helical domains, and when a v-SNARE interacts with a t-SNARE, the helical domains of one wrap around the helical domains of the other to form a stable four-helix bundle. The resulting *trans-SNARE complexes* lock the two membranes together.

SNAREs have been best characterized in neurons, where they mediate the docking and fusion of synaptic vesicles at the nerve terminal's plasma membrane in the process of neurotransmitter release. The bacteria that cause tetanus and botulism secrete powerful proteolytic neurotoxins that enter specific neurons and cleave SNARE proteins in the nerve terminals. In this way, the toxins block synaptic transmissions, often leading to death.

It is thought that the trans-SNARE complexes catalyze membrane fusion by using the energy that is freed when the interacting helices wrap around each other to pull the membrane faces together, simultaneously squeezing out water molecules from the interface (**Figure 13–17**). When liposomes containing purified v-SNAREs are mixed with liposomes containing complementary t-SNAREs, their membranes fuse, albeit slowly. In the cell, other proteins recruited to the fusion site, presumably Rab effectors, cooperate with SNAREs to accelerate fusion.

Fusion does not always follow immediately after v-SNAREs and t-SNAREs pair. As we discuss later, in the process of regulated exocytosis, fusion is delayed until secretion is triggered by a specific extracellular signal. In this case, a localized influx of Ca^{2+} triggers fusion, presumably by releasing inhibitory proteins that prevent the complete zipping-up of the trans-SNARE complexes.

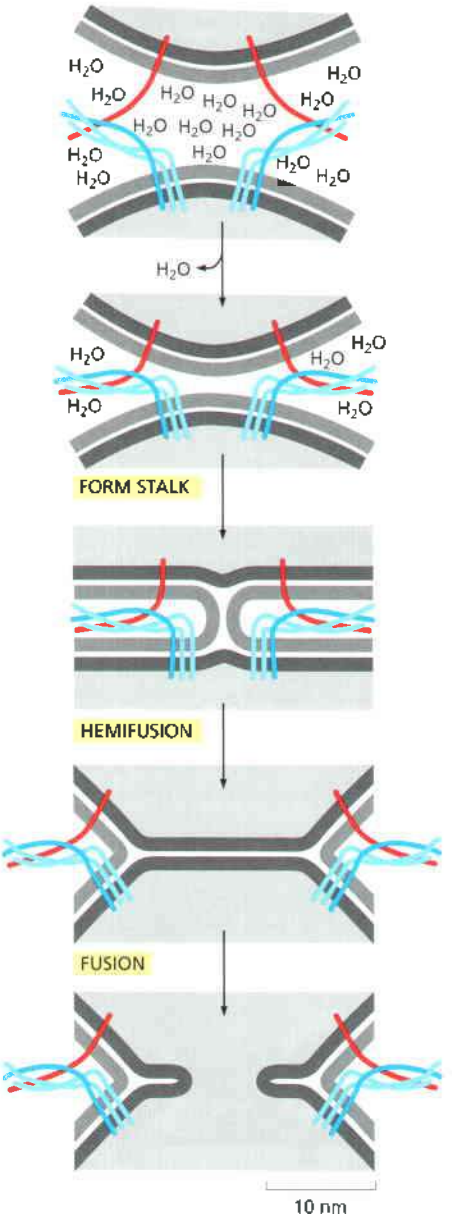


Figure 13–17 A model for how SNARE proteins may catalyze membrane fusion. Bilayer fusion occurs in multiple steps. A tight pairing between v- and t-SNAREs forces lipid bilayers into close apposition and expels water molecules from the interface. Lipid molecules in the two interacting leaflets of the bilayers then flow between the membranes to form a connecting stalk. Lipids of the other two leaflets then contact each other, forming a new bilayer, which widens the fusion zone (*hemifusion*, or half-fusion). Rupture of the new bilayer completes the fusion reaction.

Rab proteins, which can regulate the availability of SNARE proteins, exert an additional layer of control. t-SNAREs in target membranes are often associated with inhibitory proteins that must be released before the t-SNARE can function. Rab proteins and their effectors trigger the release of such SNARE inhibitory proteins. In this way, SNARE proteins are concentrated and activated in the correct location on the membrane, where tethering proteins capture incoming vesicles. Rab proteins thus speed up the process by which appropriate SNARE proteins in two membranes find each other.

For vesicular transport to operate normally, transport vesicles must incorporate the appropriate SNARE and Rab proteins. Not surprisingly, therefore, many transport vesicles will form only if they incorporate the appropriate complement of SNARE and Rab proteins in their membrane. How this crucial control process operates during vesicle budding remains a mystery.

Interacting SNAREs Need to Be Pried Apart Before They Can Function Again

Most SNARE proteins in cells have already participated in multiple rounds of vesicular transport and are sometimes present in a membrane as stable complexes with partner SNAREs. The complexes have to disassemble before the SNAREs can mediate new rounds of transport. A crucial protein called **NSF** cycles between membranes and the cytosol and catalyzes the disassembly process. It is an ATPase that uses the energy of ATP hydrolysis to unravel the intimate interactions between the helical domains of paired SNARE proteins (**Figure 13–18**). The requirement for NSF-mediated reactivation of SNAREs by SNARE complex disassembly helps prevent membranes from fusing indiscriminately: if the t-SNAREs in a target membrane were always active, any membrane containing an appropriate v-SNARE might fuse whenever the two membranes made contact. It is not known how the activity of NSF is controlled so that the SNARE machinery is activated at the right time and place, but Rab effectors are likely candidates to play a part in this process.

Viral Fusion Proteins and SNAREs May Use Similar Fusion Mechanisms

Membrane fusion is important in other processes beside vesicular transport. The plasma membranes of a sperm and an egg fuse during fertilization (discussed in Chapter 21), and myoblasts fuse with one another during the development of multinucleate muscle fibers (discussed in Chapter 22). Likewise, mitochondria fuse and fragment in a dynamic way (discussed in Chapter 14). All cell membrane fusions require special proteins and are tightly regulated to ensure that only appropriate membranes fuse. The controls are crucial for maintaining both the identity of cells and the individuality of each type of intracellular compartment.

The membrane fusions catalyzed by viral fusion proteins are the best understood. These proteins have a crucial role in permitting the entry of enveloped viruses (which have a lipid-bilayer-based membrane coat) into the cells that they infect (discussed in Chapters 5 and 24). For example, viruses such as the human immunodeficiency virus (HIV), which causes AIDS, bind to cell-surface

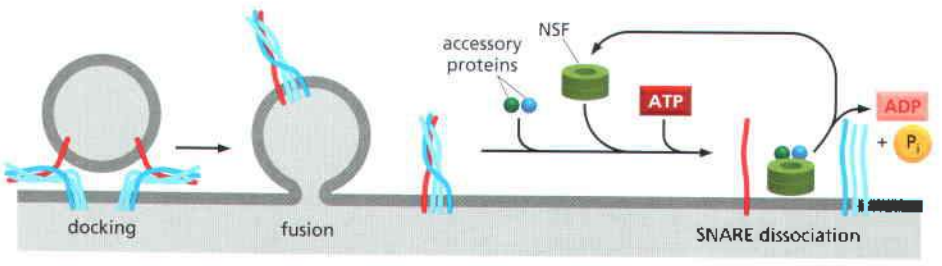


Figure 13–18 **Dissociation of SNARE pairs by NSF after a membrane fusion cycle.** After a v-SNARE and t-SNARE have mediated the fusion of a transport vesicle on a target membrane, the NSF binds to the SNARE complex and, with the help of two accessory proteins, hydrolyzes ATP to pry the SNAREs apart.

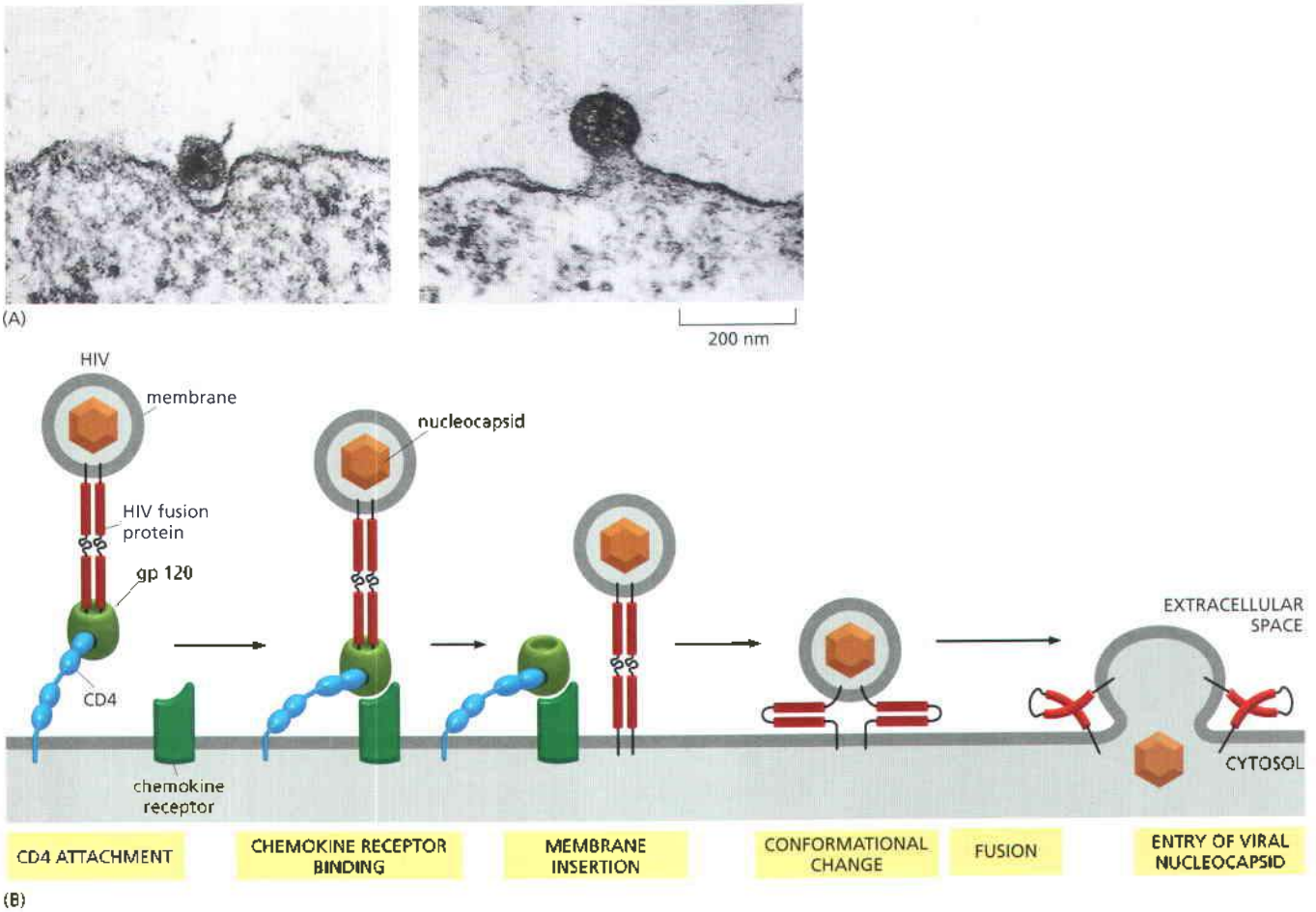


Figure 13–19 The entry of enveloped viruses into cells. (A) Electron micrographs showing how HIV enters a cell by fusing its membrane with the plasma membrane of the cell. (B) A model for the fusion process. HIV binds first to the CD4 protein on the surface of the target cell. The viral gp120 protein, which is bound to the HIV fusion protein, mediates this interaction. A second cell-surface protein on the target cell, which normally serves as a receptor for chemokines (discussed in Chapters 24 and 25), now interacts with gp120. This interaction releases the HIV fusion protein from gp120, allowing the previously buried hydrophobic fusion peptide to insert into the plasma membrane of the target cell. The fusion protein, which is a trimer (not shown), thus becomes transiently anchored as an integral membrane protein in the two opposing membranes. The fusion protein then spontaneously rearranges, collapsing into a tightly packed six-helix bundle. The energy released by this rearrangement in multiple copies of the fusion protein pulls the two membranes together, overcoming the high activation energy barrier that normally prevents membrane fusion. Thus, like a mouse trap, the HIV fusion protein contains a reservoir of potential energy, which is released and harnessed to do mechanical work. (A, from B.S. Stein et al., *Cell* 49:659–668, 1987. With permission from Elsevier; B, adapted from a drawing by Wayne Hendrickson.)

receptors and then fuse with the plasma membrane of the target cell (**Figure 13–19**). This fusion event allows the viral nucleic acid inside the nucleocapsid to enter the cytosol, where it replicates. Other viruses, such as the influenza virus, first enter the cell by receptor-mediated endocytosis (discussed later) and are delivered to endosomes; the low pH in endosomes activates a fusion protein in the viral envelope that catalyzes the fusion of the viral and endosomal membranes, releasing the viral nucleic acid into the cytosol.

The three-dimensional structures of the fusion proteins of HIV and influenza virus provide valuable insights into the molecular mechanism of the membrane fusion catalyzed by these proteins. Exposure of the HIV fusion protein to receptors on the target cell membrane, or exposure of the influenza fusion protein to low pH, uncovers previously buried hydrophobic regions. These regions, called *fusion peptides*, then insert directly into the lipid bilayer of the target cell membrane. In this way, the fusion proteins transiently become integral membrane proteins in two separate lipid bilayers. Structural rearrangements in the fusion proteins then bring the two lipid bilayers into very close apposition and destabilize them so that the bilayers fuse (see Figure 13–19). Thus, viral fusion proteins and SNAREs promote lipid bilayer fusion in similar ways.

Summary

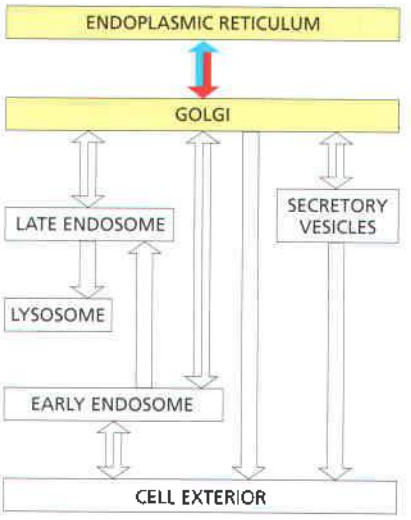
Directed and selective transport of particular membrane components from one membrane-enclosed compartment of a eucaryotic cell to another maintains the differences between those compartments. Transport vesicles, which can be spherical, tubular, or irregularly shaped, bud from specialized coated regions of the donor membrane. The assembly of the coat helps to collect specific membrane and soluble cargo molecules for transport and to drive the formation of the vesicle.

There are various types of coated vesicles. The best characterized are clathrin-coated vesicles, which mediate transport from the plasma membrane and the trans *Golgi network, and COPI- and COPII-coated vesicles, which mediate transport between Golgi cisternae and between the ER and the Golgi apparatus, respectively. In clathrin-coated vesicles, adaptor proteins link the clathrin to the vesicle membrane and also trap specific cargo molecules for packaging into the vesicle. The coat is shed rapidly after budding, enabling the vesicle to fuse with its appropriate target membrane.*

Local synthesis of phosphoinositides creates binding sites that trigger coat assembly and vesicle budding. In addition, monomeric GTPases help regulate various steps in vesicular transport, including both vesicle budding and docking. The coat-recruitment GTPases, including Sar1 and the Arf proteins, regulate coat assembly and disassembly. A large family of Rab proteins functions as vesicle targeting GTPases. Rab proteins are recruited to transport vesicles and target membranes. The assembly and disassembly of Rab proteins and their effectors in specialized membrane domains are dynamically controlled by GTP binding and hydrolysis. Active Rab proteins recruit Rab effectors, such as motor proteins, which transport vesicles on actin filaments or microtubules, and filamentous tethering proteins, which help ensure that the vesicles deliver their contents only to the appropriate membrane-enclosed compartment. Complementary v-SNARE proteins on transport vesicles and t-SNARE proteins on the target membrane form stable trans-SNARE complexes, which force the two membranes into close apposition so that their lipid bilayers can fuse.

TRANSPORT FROM THE ER THROUGH THE GOLGI APPARATUS

As discussed in Chapter 12, newly synthesized proteins cross the ER membrane from the cytosol to enter the biosynthetic–secretory pathway. During their subsequent transport, from the ER to the Golgi apparatus and from the Golgi apparatus to the cell surface and elsewhere, these proteins are successively modified as they pass through a series of compartments. Transfer from one compartment to the next involves a delicate balance between forward and backward (retrieval) transport pathways. Some transport vesicles select cargo molecules and move them to the next compartment in the pathway, while others retrieve escaped proteins and return them to a previous compartment where they normally function. Thus, the pathway from the ER to the cell surface consists of many sorting steps, which continuously select membrane and soluble lumenal proteins for packaging and transport—in vesicles or organelle fragments that bud from the ER and Golgi apparatus.



In this section we focus mainly on the **Golgi apparatus** (also called the **Golgi complex**). It is a major site of carbohydrate synthesis, as well as a sorting and dispatching station for products of the ER. The cell makes many polysaccharides in the Golgi apparatus, including the pectin and hemicellulose of the cell wall in plants and most of the glycosaminoglycans of the extracellular matrix in animals (discussed in Chapter 19). The Golgi apparatus also lies on the exit route from the ER, and a large proportion of the carbohydrates that it makes are attached as oligosaccharide side chains to the many proteins and lipids that the ER sends to it. A subset of these oligosaccharide groups serve as tags to direct specific proteins into vesicles that then transport them to lysosomes. But most proteins and lipids, once they have acquired their appropriate oligosaccharides in the Golgi apparatus, are recognized in other ways for targeting into the transport vesicles going to other destinations.

Proteins Leave the ER in COPII-Coated Transport Vesicles

To initiate their journey along the biosynthetic–secretory pathway, proteins that have entered the ER and are destined for the Golgi apparatus or beyond are first packaged into small COPII-coated transport vesicles. These vesicles bud from specialized regions of the ER called *ER exit sites*, whose membrane lacks bound ribosomes. Most animal cells have ER exit sites dispersed throughout the ER network.

Originally, it was thought that all proteins that are not tethered in the ER enter transport vesicles by default. It is now clear, however, that entry into vesicles that leave the ER is usually a selective process. Many membrane proteins are actively recruited into such vesicles, where they become concentrated. It is thought that these cargo proteins display exit (transport) signals on their cytosolic surface that components of the COPII coat recognize (**Figure 13–20**); these coat components act as cargo receptors and are recycled back to the ER after they have delivered their cargo to the Golgi apparatus. Soluble cargo proteins in the ER lumen, by contrast, have exit signals that attach them to transmembrane cargo receptors, which in turn bind through exit signals in their cytoplasmic tails to components of the COPII coat. At a lower rate, proteins without exit signals can also enter transport vesicles, so that even proteins that normally function in the ER (so-called *ER resident proteins)* slowly leak out of the ER and are delivered to the Golgi apparatus. Similarly, secretory proteins that are made in high concentrations may leave the ER without the help of exit signals or cargo receptors.

The exit signals that direct soluble proteins out of the ER for transport to the Golgi apparatus and beyond are not well understood. Some transmembrane proteins that serve as cargo receptors for packaging some secretory proteins into COPII-coated vesicles are lectins that bind to oligosaccharides. The ERGIC53 lectin, for example, binds to mannose and is thought to recognize this sugar on two secreted blood-clotting factors (Factor V and Factor VIII), thereby packaging the proteins into transport vesicles in the ER. ERGIC53's role in protein transport was identified because humans who lack it owing to an inherited mutation have lowered serum levels of Factors V and VIII, and they therefore bleed excessively.

Only Proteins That Are Properly Folded and Assembled Can Leave the ER

To exit from the ER, proteins must be properly folded and, if they are subunits of multimeric protein complexes, they may need to be completely assembled. Those that are misfolded or incompletely assembled remain in the ER, where they are bound to chaperone proteins (discussed in Chapter 6), such as *BiP* or *calnexin*. The chaperones may cover up the exit signals or somehow anchor the proteins in

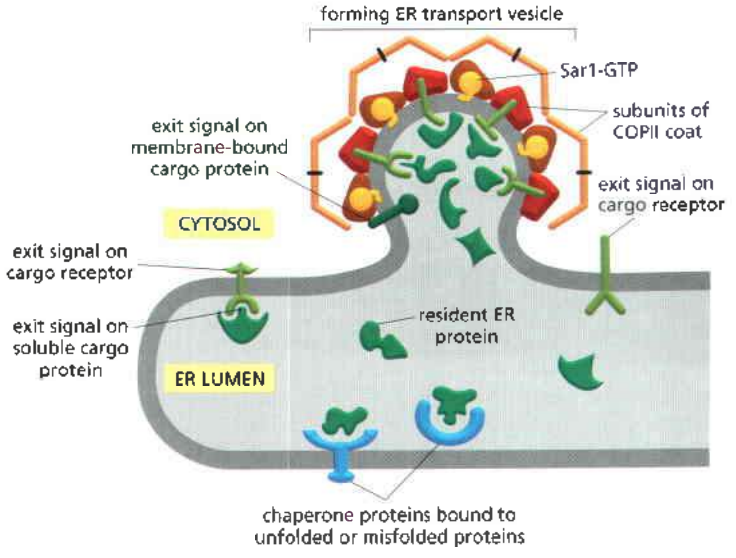


Figure 13–20 The recruitment of cargo molecules into ER transport vesicles. By binding directly or indirectly to the COPII coat, membrane and soluble cargo proteins, respectively, become concentrated in the transport vesicles as they leave the ER. Membrane proteins are packaged into budding transport vesicles through interactions of exit signals on their cytosolic tails with the COPII coat. Some of the membrane proteins that the coat traps function as cargo receptors, binding soluble proteins in the lumen and helping to package them into vesicles. A typical 50-nm transport vesicle contains about 200 membrane proteins, which can be of many different types. As indicated, unfolded or incompletely assembled proteins are bound to chaperones and retained in the ER compartment.

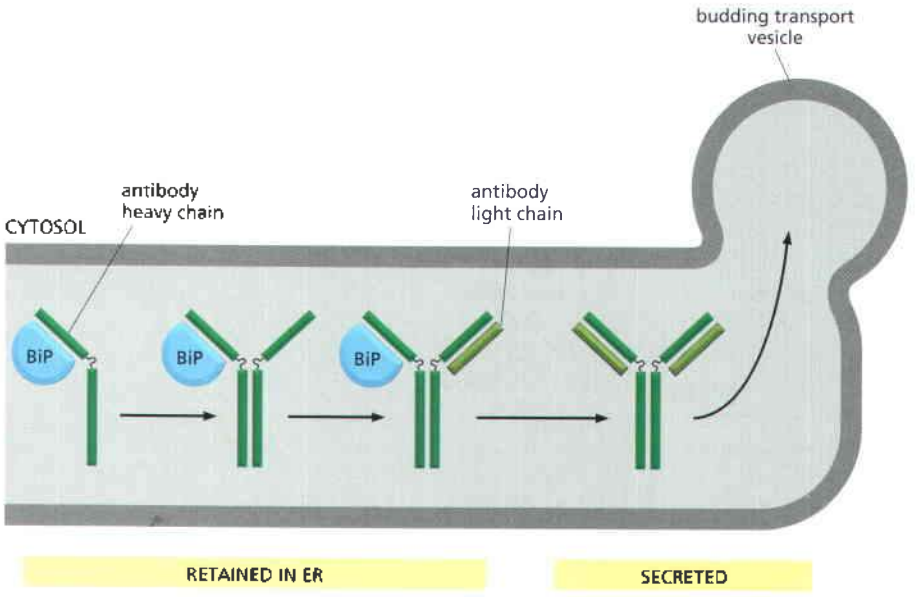


Figure 13–21 Retention of incompletely assembled antibody molecules in the ER. Antibodies are made up of two heavy and two light chains (discussed in Chapter 25), which assemble in the ER. The chaperone BiP is thought to bind to all incompletely assembled antibody molecules and to cover up an exit signal. Thus, only completely assembled antibodies leave the ER and are secreted.

the ER (Figure 13–21). Such failed proteins are eventually transported back into the cytosol, where they are degraded by proteasomes (discussed in Chapters 6 and 12). This quality-control step prevents the onward transport of misfolded or misassembled proteins that could potentially interfere with the functions of normal proteins. There is a surprising amount of corrective action. More than 90% of the newly synthesized subunits of the T cell receptor (discussed in Chapter 25) and of the acetylcholine receptor (discussed in Chapter 11), for example, are normally degraded without ever reaching the cell surface where they function. Thus, cells must make a large excess of many protein molecules to produce a select few that fold, assemble, and function properly.

The process of continual degradation of a portion of ER proteins also provides an early warning system to alert the immune system when a virus infects cells. Using specialized ABC-type transporters, the ER imports peptide fragments of viral proteins produced by proteases in the proteasome. The foreign peptides are loaded onto class I MHC proteins in the ER lumen and then transported to the cell surface. T lymphocytes then recognize the peptides as non-self antigens and kill the infected cells (discussed in Chapter 25).

Sometimes, however, there are drawbacks to the stringent quality-control mechanism. The predominant mutations that cause cystic fibrosis, a common inherited disease, result in the production of a slightly misfolded form of a plasma membrane protein important for Cl^- transport. Although the mutant protein would function perfectly normally if it reached the plasma membrane, it remains in the ER. This devastating disease thus results not because the mutation inactivates the protein but because the active protein is discarded before it reaches the plasma membrane.

Vesicular Tubular Clusters Mediate Transport from the ER to the Golgi Apparatus

After transport vesicles have budded from ER exit sites and have shed their coat, they begin to fuse with one another. This fusion of membranes from the same compartment is called *homotypic fusion*, to distinguish it from *heterotypic fusion*, in which a membrane from one compartment fuses with the membrane of a different compartment. As with heterotypic fusion, homotypic fusion requires a set of matching SNAREs. In this case, however, the interaction is symmetrical, with both membranes contributing v-SNAREs and t-SNAREs (Figure 13–22).

The structures formed when ER-derived vesicles fuse with one another are called *vesicular tubular clusters*, because they have a convoluted appearance in

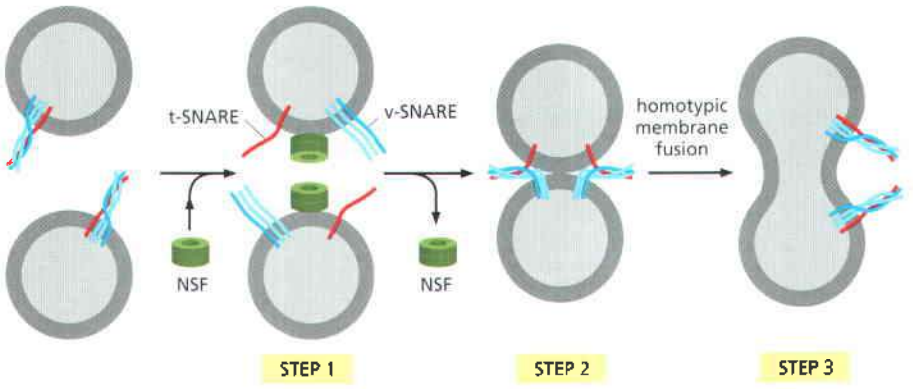


Figure 13–22 Homotypic membrane fusion. <AAAA> In step 1, NSF pries apart identical pairs of v-SNAREs and t-SNAREs in both membranes (see Figure 13–18). In steps 2 and 3, the separated matching SNAREs on adjacent identical membranes interact, which leads to membrane fusion and the formation of one continuous compartment called a vesicular tubular cluster. Subsequently, the compartment grows by further homotypic fusion with vesicles from the same kind of membrane, displaying matching SNAREs. Homotypic fusion is not restricted to the formation of vesicular tubular clusters; in a similar process, endosomes fuse to generate larger endosomes. Rab proteins help regulate the extent of homotypic fusion and hence the size of the compartments in a cell (not shown).

the electron microscope (**Figure 13–23A**). These clusters constitute a new compartment that is separate from the ER and lacks many of the proteins that function in the ER. They are generated continually and function as transport containers that bring material from the ER to the Golgi apparatus. The clusters are relatively short-lived because they move quickly along microtubules to the Golgi apparatus, with which they fuse to deliver their contents (Figure 13–23B).

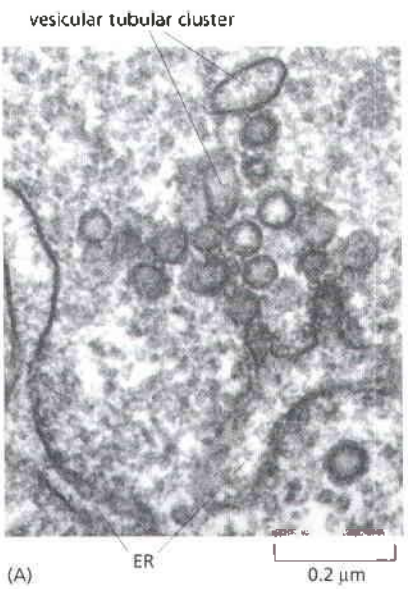
As soon as vesicular tubular clusters form, they begin to bud off transport vesicles of their own. Unlike the COPII-coated vesicles that bud from the ER, these vesicles are COPI-coated. They carry back to the ER resident proteins that have escaped, as well as proteins such as cargo receptors that participated in the ER budding reaction and are being returned. This retrieval process demonstrates the exquisite control mechanisms that regulate coat assembly reactions. The COPI coat assembly begins only seconds after the COPII coats have been shed. It remains a mystery how this switch in coat assembly is controlled.

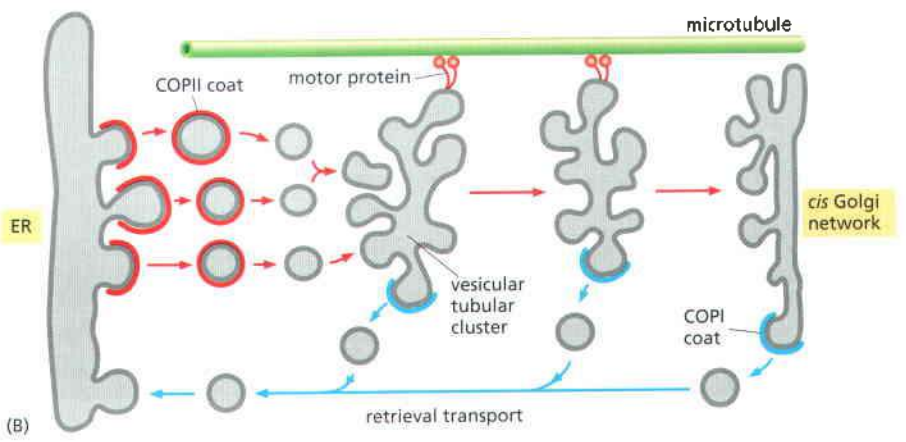
The *retrieval* (or *retrograde*) *transport* continues as the vesicular tubular clusters move towards the Golgi apparatus. Thus, the clusters continuously mature, gradually changing their composition as selected proteins are returned to the ER. A similar retrieval process continues from the Golgi apparatus, after the vesicular tubular clusters have delivered their cargo.

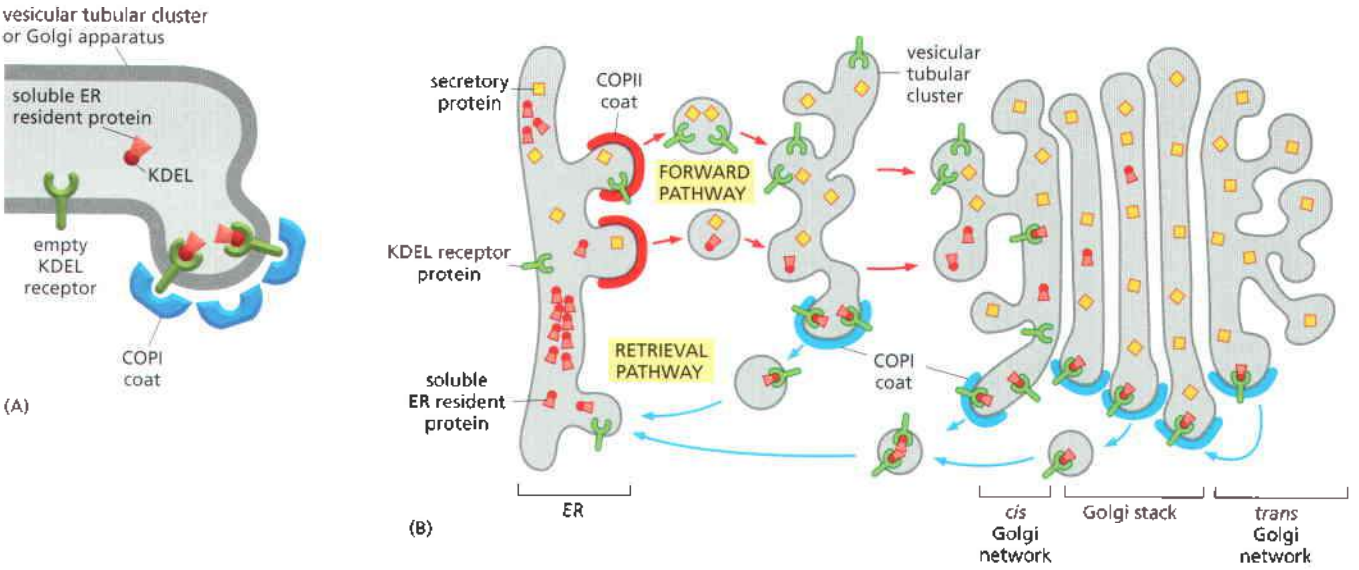
The Retrieval Pathway to the ER Uses Sorting Signals

The retrieval pathway for returning escaped proteins back to the ER depends on *ER retrieval signals*. Resident ER membrane proteins, for example, contain signals that bind directly to COPI coats and are thus packaged into COPI-coated

Figure 13–23 Vesicular tubular clusters. (A) An electron micrograph section of vesicular tubular clusters forming from the ER membrane. Many of the vesicle-like structures seen in the micrograph are cross sections of tubules that extend above and below the plane of this thin section and are interconnected. (B) Vesicular tubular clusters move along microtubules to carry proteins from the ER to the Golgi apparatus. COPI coats mediate the budding of vesicles that return to the ER from these clusters. As indicated, the coats quickly disassemble after the vesicles have formed. (A, courtesy of William Balch.)







transport vesicles for retrograde delivery to the ER. The best-characterized retrieval signal of this type consists of two lysines, followed by any two other amino acids, at the extreme C-terminal end of the ER membrane protein. It is called a *KKXX sequence*, based on the single-letter amino acid code.

Soluble ER resident proteins, such as BiP, also contain a short retrieval signal at their C-terminal end, but it is different: it consists of a Lys-Asp-Glu-Leu or a similar sequence. If this signal (called the *KDEL sequence*) is removed from BiP by genetic engineering, the protein is slowly secreted from the cell. If the signal is transferred to a protein that is normally secreted, the protein is now efficiently returned to the ER, where it accumulates.

Unlike the retrieval signals on ER membrane proteins, which can interact directly with the COPI coat, soluble ER resident proteins must bind to specialized receptor proteins such as the *KDEL receptor*—a multipass transmembrane protein that binds to the KDEL sequence and packages any protein displaying it into COPI-coated retrograde transport vesicles (**Figure 13–24**). To accomplish this task, the KDEL receptor itself must cycle between the ER and the Golgi apparatus, and its affinity for the KDEL sequence must differ in these two compartments. The receptor must have a high affinity for the KDEL sequence in vesicular tubular clusters and the Golgi apparatus, so as to capture escaped, soluble ER resident proteins that are present there at low concentration. It must have a low affinity for the KDEL sequence in the ER, however, to unload its cargo in spite of the very high concentration of KDEL-containing resident proteins in the ER.

How does the affinity of the KDEL receptor change depending on the compartment in which it resides? The answer is not known, but it may be related to the different ionic conditions and pH in the different compartments, which are regulated by ion transporters in the compartment membrane. As we discuss later, pH-sensitive protein–protein interactions form the basis for many of the sorting steps in the cell.

Most membrane proteins that function at the interface between the ER and Golgi apparatus, including v- and t-SNAREs and some cargo receptors, enter the retrieval pathway back to the ER. Whereas the recycling of some of these proteins is mediated by signals, as just described, for others no specific signal seems to be required. Thus, while retrieval signals increase the efficiency of the retrieval process, some proteins randomly enter budding vesicles destined for the ER and are returned to the ER at a slower rate. Many Golgi enzymes cycle constantly between the ER and the Golgi, but their rate of return to the ER is slow enough for most of the protein to be found in the Golgi apparatus.

Figure 13–24 A model for the retrieval of soluble ER resident proteins. ER resident proteins that escape from the ER are returned by vesicular transport. (A) The KDEL receptor present in vesicular tubular clusters and the Golgi apparatus captures the soluble ER resident proteins and carries them in COPI-coated transport vesicles back to the ER. Upon binding its ligands in this environment, the KDEL receptor may change conformation, so as to facilitate its recruitment into budding COPI-coated vesicles. (B) The retrieval of ER proteins begins in vesicular tubular clusters and continues from all parts of the Golgi apparatus. In the environment of the ER, the ER resident proteins dissociate from the KDEL receptor, which is then returned to the Golgi apparatus for reuse.

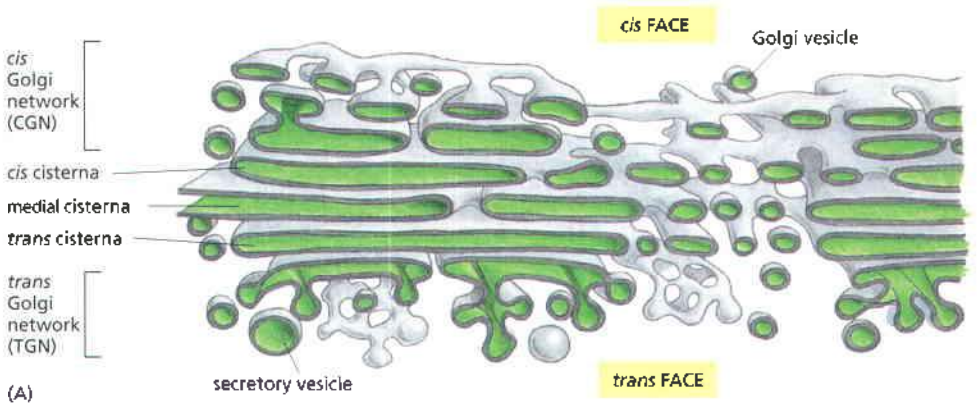
Many Proteins Are Selectively Retained in the Compartments in Which They Function

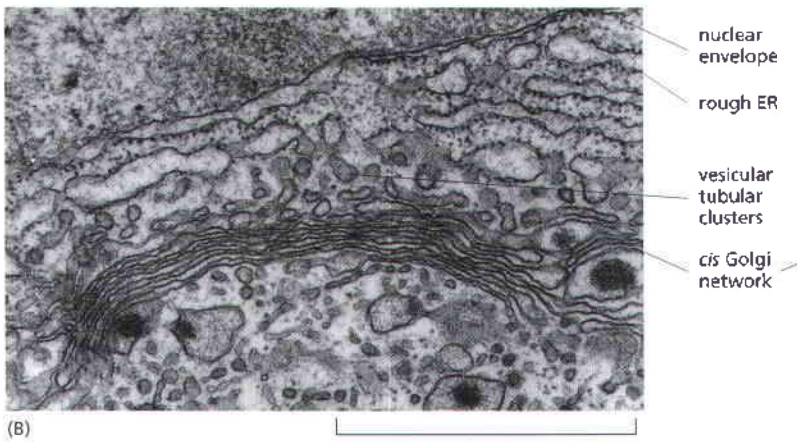
The KDEL retrieval pathway only partly explains how ER resident proteins are maintained in the ER. As expected, cells that express genetically modified ER resident proteins, from which the KDEL sequence has been experimentally removed, secrete these proteins. But the rate of secretion is much slower than for a normal secretory protein. It seems that a mechanism that is independent of their KDEL signal anchors ER resident proteins and that only those proteins that escape this retention mechanism are captured and returned via the KDEL receptor. A suggested retention mechanism is that ER resident proteins bind to one another, thus forming complexes that are too big to enter transport vesicles efficiently. Because ER resident proteins are present in the ER at very high concentrations (estimated to be millimolar), relatively low-affinity interactions would suffice to tie up most of the proteins in such complexes.

Aggregation of proteins that function in the same compartment—called *kin recognition*—is a general mechanism that compartments use to organize and retain their resident proteins. Golgi enzymes that function together, for example, also bind to each other and are thereby restrained from entering transport vesicles leaving the Golgi apparatus.

The Golgi Apparatus Consists of an Ordered Series of Compartments

Because of its large and regular structure, the Golgi apparatus was one of the first organelles described by early light microscopists. It consists of a collection of flattened, membrane-enclosed compartments, called *cisternae*, that somewhat resemble a stack of pita breads. Each Golgi stack usually consists of four to six cisternae (**Figure 13–25**), although some unicellular flagellates can have up to





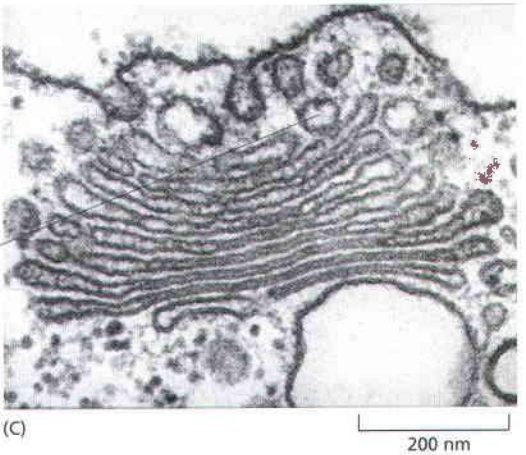


Figure 13–25 The Golgi apparatus. (A) Three-dimensional reconstruction from electron micrographs of the Golgi apparatus in a secretory animal cell. The *cis* face of the Golgi stack is that closest to the ER. (B) A thin-section electron micrograph emphasizing the transitional zone between the ER and the Golgi apparatus in an animal cell. (C) An electron micrograph of a Golgi apparatus in a plant cell (the green alga *Chlamydomonas*) seen in cross section. In plant cells, the Golgi apparatus is generally more distinct and more clearly separated from other intracellular membranes than in animal cells. (A, redrawn from A. Rambourg and Y. Clermont, *Eur. J. Cell Biol.* 51:189–200, 1990. With permission from Wissenschaftliche Verlagsgesellschaft; B, courtesy of Brij J. Gupta; C, courtesy of George Palade.)

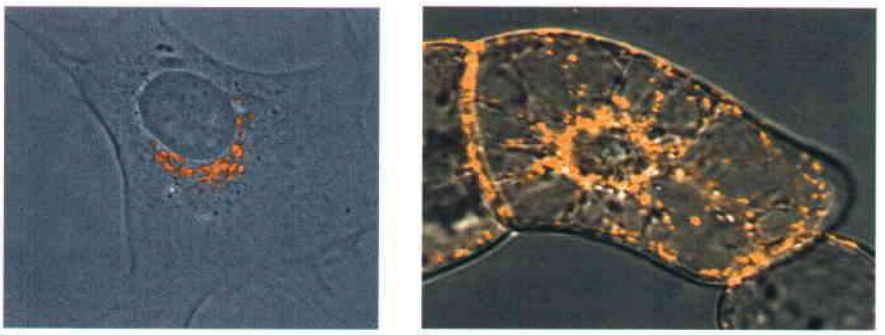


Figure 13–26 Localization of the Golgi apparatus in animal and plant cells. (A) The Golgi apparatus in a cultured fibroblast stained with a fluorescent antibody that recognizes a Golgi resident protein (*red*). The Golgi apparatus is polarized, facing the direction in which the cell was crawling before fixation. (B) The Golgi apparatus in a plant cell that is expressing a fusion protein consisting of a resident Golgi enzyme fused to green fluorescent protein. The bright *orange* spots (false color) are Golgi stacks. (A, courtesy of John Henley and Mark McNiven; B, courtesy of Chris Hawes.)

60. In animal cells, tubular connections between corresponding cisternae link many stacks, thus forming a single complex, which is usually located near the cell nucleus and close to the centrosome (Figure 13–26A). This localization depends on microtubules. If microtubules are experimentally depolymerized, the Golgi apparatus reorganizes into individual stacks that are found throughout the cytoplasm, adjacent to ER exit sites. Some cells, including most plant cells, have hundreds of individual Golgi stacks dispersed throughout the cytoplasm (Figure 13–26B).

During their passage through the Golgi apparatus, transported molecules undergo an ordered series of covalent modifications. Each Golgi stack has two distinct faces: a ***cis* face** (or entry face) and a ***trans* face** (or exit face). Both *cis* and *trans* faces are closely associated with special compartments, each composed of a network of interconnected tubular and cisternal structures: the ***cis* Golgi network (CGN)** and the ***trans* Golgi network (TGN)**, respectively. The CGN is a collection of fused vesicular tubular clusters arriving from the ER. Proteins and lipids enter the *cis* Golgi network and exit from the *trans* Golgi network bound for the cell surface or another compartment. Both networks are important for protein sorting: proteins entering the CGN can either move onward in the Golgi apparatus or be returned to the ER. Similarly, proteins exiting from the TGN move onward and are sorted according to their next destination: lysosomes, secretory vesicles, or the cell surface. They also can be returned to an earlier compartment.

As described in Chapter 12, a single species of *N-linked oligosaccharide* is attached *en bloc* to many proteins in the ER and then trimmed while the protein is still in the ER. The oligosaccharide intermediates created by the trimming reactions serve to help proteins fold and to help transport misfolded proteins to the cytosol for degradation. Thus, they play an important role in controlling the quality of proteins exiting from the ER. Once these ER functions have been fulfilled, the cell is free to redesign the oligosaccharides for new functions. This happens in the Golgi apparatus, which generates the heterogeneous oligosaccharide structures seen in mature proteins. Upon arrival in the CGN, proteins pass through the *cis* Golgi network, before entering the first of the Golgi processing compartments (the *cis* Golgi cisternae). They then move to the next compartment (the medial cisternae) and finally to the *trans* cisternae, where glycosylation is completed. The lumen of the *trans* cisternae is thought to be continuous with the TGN, the place where proteins are segregated into different transport packages and dispatched to their final destinations.

The oligosaccharide processing steps occur in an organized sequence in the Golgi stack, with each cisterna containing a characteristic abundance of processing enzymes. Proteins are modified in successive stages as they move from cisterna to cisterna across the stack, so that the stack forms a multistage processing unit. This compartmentalization might seem unnecessary, since each oligosaccharide processing enzyme can accept a glycoprotein as a substrate only after it has been properly processed by the preceding enzyme. Nonetheless, it is clear that processing occurs in a spatial as well as a biochemical sequence: enzymes catalyzing early processing steps are concentrated in the cisternae toward the *cis* face of the Golgi stack, whereas enzymes catalyzing later processing steps are concentrated in the cisternae toward the *trans* face.

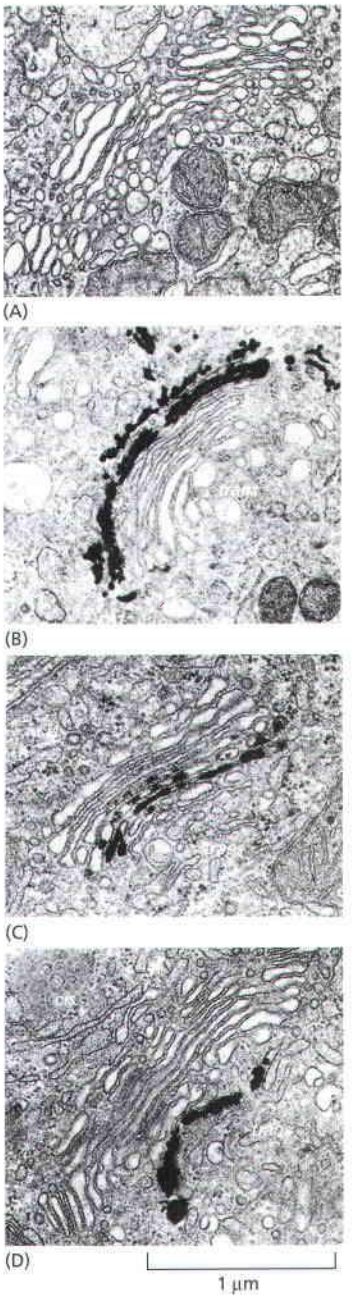


Figure 13–27 Molecular compartmentalization of the Golgi apparatus. A series of electron micrographs shows the Golgi apparatus (A) unstained, (B) stained with osmium, which is preferentially reduced by the cisternae of the *cis* compartment, and (C and D) stained to reveal the location of specific enzymes. Nucleoside diphosphatase is found in the *trans* Golgi cisternae (C), while acid phosphatase is found in the *trans* Golgi network (D). Note that usually more than one cisterna is stained. The enzymes are therefore thought to be highly enriched rather than precisely localized to a specific cisterna. (Courtesy of Daniel S. Friend.)

Investigators discovered the functional differences between the *cis*, medial, and *trans* subdivisions of the Golgi apparatus by localizing the enzymes involved in processing *N*-linked oligosaccharides in distinct regions of the organelle, both by physical fractionation of the organelle and by labeling the enzymes in electron microscope sections with antibodies. The removal of mannose residues and the addition of *N*-acetylglucosamine, for example, were shown to occur in the medial compartment, while the addition of galactose and sialic acid was found to occur in the *trans* compartment and the *trans* Golgi network (**Figure 13–27**). **Figure 13–28** summarizes the functional compartmentalization of the Golgi apparatus.

The Golgi apparatus is especially prominent in cells that are specialized for secretion of glycoproteins, such as the goblet cells of the intestinal epithelium, which secrete large amounts of polysaccharide-rich mucus into the gut (**Figure 13–29**). In such cells, unusually large vesicles are found on the *trans* side of the Golgi apparatus, which faces the plasma membrane domain where secretion occurs.

Oligosaccharide Chains Are Processed in the Golgi Apparatus

Whereas the ER lumen is full of soluble lumenal resident proteins and enzymes, the resident proteins in the Golgi apparatus are all membrane bound. The enzymatic reactions in the Golgi apparatus seem to be carried out entirely on its membrane surface. All of the Golgi glycosidases and glycosyl transferases are single-pass transmembrane proteins, many of which are organized in multienzyme complexes. Thus, the two synthetic organelles in the biosynthetic-secretory pathway, the ER and the Golgi apparatus, are organized in fundamentally different ways.

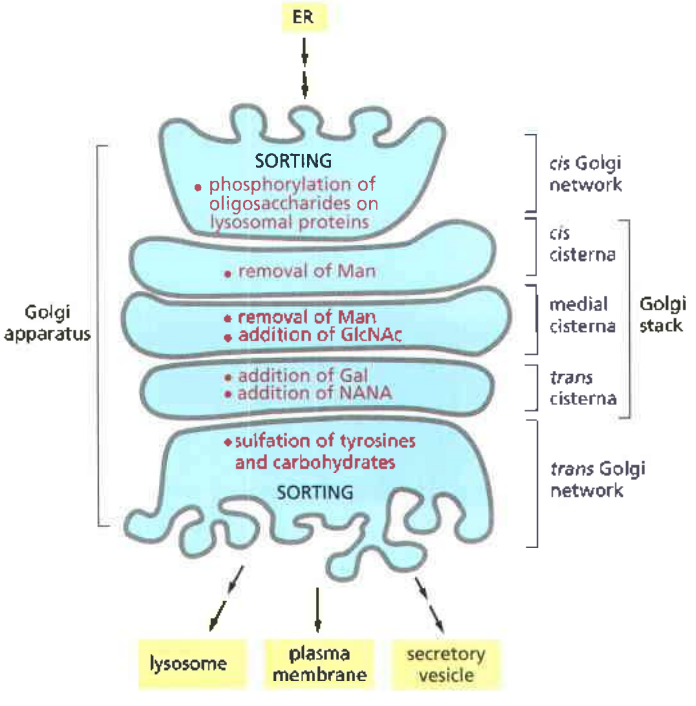


Figure 13–28 Oligosaccharide processing in Golgi compartments. The localization of each processing step shown was determined by a combination of techniques, including biochemical subfractionation of the Golgi apparatus membranes and electron microscopy after staining with antibodies specific for some of the processing enzymes. Processing enzymes are not restricted to a particular cisterna; instead, their distribution is graded across the stack—such that early-acting enzymes are present mostly in the *cis* Golgi cisternae and later-acting enzymes are mostly in the *trans* Golgi cisternae.

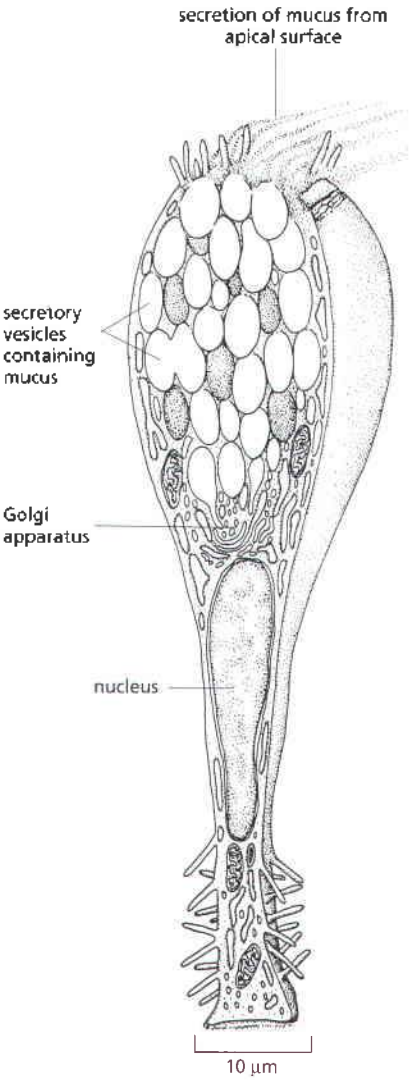


Figure 13–29 A goblet cell of the small intestine. This cell is specialized for secreting mucus, a mixture of glycoproteins and proteoglycans synthesized in the ER and Golgi apparatus. Like all epithelial cells, goblet cells are highly polarized, with the apical domain of their plasma membrane facing the lumen of the gut and the basolateral domain facing the basal lamina. The Golgi apparatus is also highly polarized, which facilitates the discharge of mucus by exocytosis at the apical domain of the plasma membrane. (After R.V. Krstic, Illustrated Encyclopedia of Human Histology. New York: Springer-Verlag, 1984. With permission from Springer-Verlag.)

Two broad classes of *N*-linked oligosaccharides, the **complex oligosaccharides** and the **high-mannose oligosaccharides**, are attached to mammalian glycoproteins (**Figure 13–30**). Sometimes, both types are attached (in different places) to the same polypeptide chain.

Complex oligosaccharides are generated when the original *N*-linked oligosaccharide added in the ER is trimmed and further sugars are added. By contrast, high-mannose oligosaccharides are trimmed but have no new sugars added to them in the Golgi apparatus. They contain just two *N*-acetylglucosamines and many mannose residues, often approaching the number originally present in the lipid-linked oligosaccharide precursor added in the ER. Complex oligosaccharides can contain more than the original two *N*-acetylglucosamines, as well as a variable number of galactose and sialic acids and, in some cases, fucose. Sialic acid is of special importance because it is the only sugar in glycoproteins that bears a negative charge. Whether a given oligosaccharide remains high-mannose or is processed depends largely on its position in the protein. If the oligosaccharide is

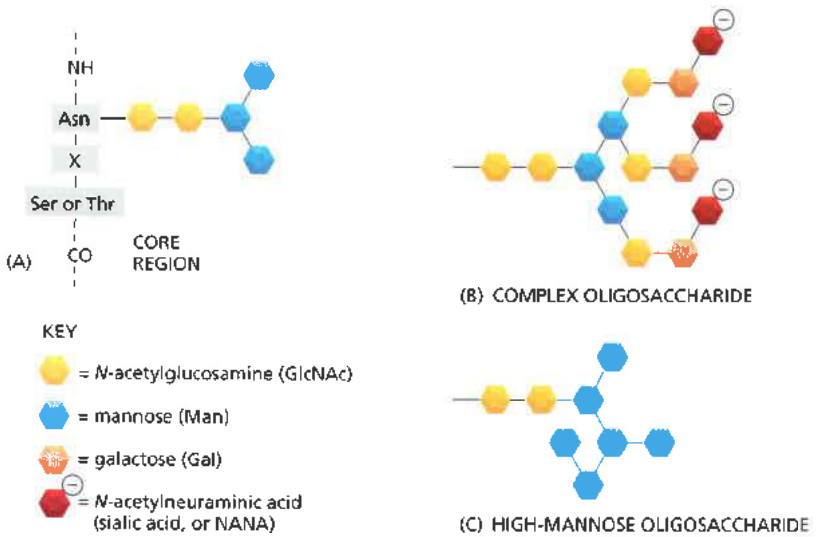


Figure 13–30 The two main classes of asparagine-linked (*N*-linked) oligosaccharides found in mature mammalian glycoproteins. (A) Both complex oligosaccharides and high-mannose oligosaccharides share a common *core region* derived from the original *N*-linked oligosaccharide added in the ER (see Figure 12–50) and typically containing two *N*-acetylglucosamines (GlcNAc) and three mannoses (Man). (B) Each complex oligosaccharide consists of a *core region*, together with a *terminal region* that contains a variable number of copies of a special trisaccharide unit *(N-acetylglucosamine–galactose–sialic acid)* linked to the core mannoses. Frequently, the terminal region is truncated and contains only GlcNAc and galactose (Gal) or just GlcNAc. In addition, a fucose residue may be added, usually to the core GlcNAc attached to the asparagine (Asn). Thus, although the steps of processing and subsequent sugar addition are rigidly ordered, complex oligosaccharides can be heterogeneous. Moreover, although the complex oligosaccharide shown has three terminal branches, two and four branches are also common, depending on the glycoprotein and the cell in which it is made. (C) High-mannose oligosaccharides are not trimmed back all the way to the core region and contain additional mannose residues. Hybrid oligosaccharides with one Man branch and one GlcNAc and Gal branch are also found (not shown).

The three amino acids indicated in (A) constitute the sequence recognized by the oligosaccharyl transferase enzyme that adds the initial oligosaccharide to the protein. Ser = serine; Thr = threonine; X = any amino acid, except proline.

accessible to the processing enzymes in the Golgi apparatus, it is likely to be converted to a complex form; if it is inaccessible because its sugars are tightly held to the protein's surface, it is likely to remain in a high-mannose form. The processing that generates complex oligosaccharide chains follows the highly ordered pathway shown in **Figure 13–31**.

Beyond these commonalities in oligosaccharide processing that are shared among most cells, the products of the carbohydrate modifications carried out in the Golgi apparatus are highly complex and have given rise to a new field called glycobiology. The human genome, for example, encodes hundreds of different Golgi glycosyl transferases, which are expressed differently from one cell type to another, resulting in a variety of glycosylated forms of a given protein or lipid in different cell types and at varying stages of differentiation, depending on the spectrum of enzymes expressed by the cell. The complexity of modifications is not limited to *N*-linked oligosaccharides but also occurs on *O-linked sugars*, as we discuss next.

Proteoglycans Are Assembled in the Golgi Apparatus

In addition to the *N*-linked oligosaccharide alterations made to proteins as they pass through the Golgi cisternae *en route* from the ER to their final destinations, many proteins are also modified in other ways. Some proteins have sugars added to the hydroxyl groups of selected serine or threonine side chains. This ***O*-linked glycosylation** (**Figure 13–32**), like the extension of *N*-linked oligosaccharide chains, is catalyzed by a series of glycosyl transferase enzymes that use the sugar nucleotides in the lumen of the Golgi apparatus to

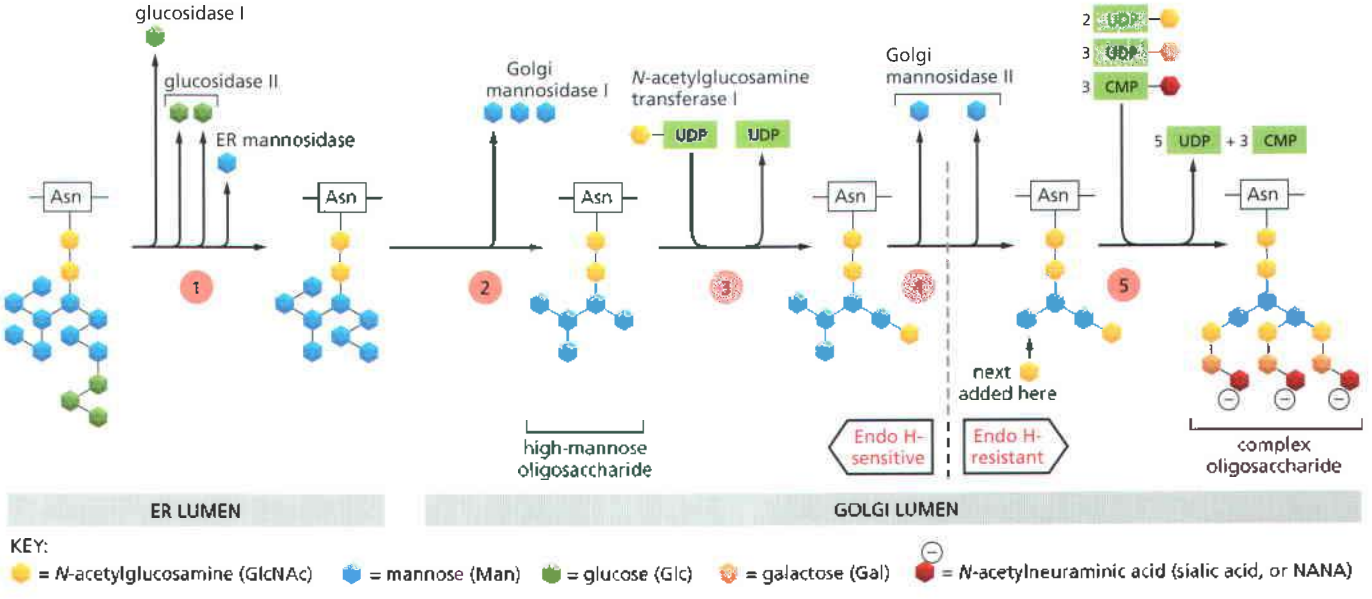


Figure 13–31 Oligosaccharide processing in the ER and the Golgi apparatus. The processing pathway is highly ordered, so that each step shown depends on the previous one. Step 1: Processing begins in the ER with the removal of the glucoses from the oligosaccharide initially transferred to the protein. Then a mannosidase in the ER membrane removes a specific mannose. Step 2: The remaining steps occur in the Golgi stack, where Golgi mannosidase I first removes three more mannoses. Step 3: *N*-acetylglucosamine transferase I then adds an *N*-acetylglucosamine. Step 4: Mannosidase II then removes two additional mannoses. This yields the final core of three mannoses that is present in a complex oligosaccharide. At this stage, the bond between the two *N*-acetylglucosamines in the core becomes resistant to attack by a highly specific endoglycosidase *(Endo H)*. Since all later structures in the pathway are also Endo H-resistant, treatment with this enzyme is widely used to distinguish complex from high-mannose oligosaccharides. Step 5: Finally, as shown in Figure 13–30, additional *N*-acetylglucosamines, galactoses, and sialic acids are added. These final steps in the synthesis of a complex oligosaccharide occur in the cisternal compartments of the Golgi apparatus. Three types of glycosyl transferase enzymes act sequentially, using sugar substrates that have been activated by linkage to the indicated nucleotide. The membranes of the Golgi cisternae contain specific carrier proteins that allow each sugar nucleotide to enter in exchange for the nucleoside phosphates that are released after the sugar is attached to the protein on the lumenal face.

Note that as a biosynthetic organelle, the Golgi apparatus differs fundamentally from the ER: all sugars are assembled inside the lumen from sugar nucleotides. By contrast, in the ER, the *N*-linked precursor oligosaccharide is assembled partly in the cytosol and partly in the lumen, and all lumenal reactions use dolichol-linked sugars as their substrates (see Figure 12–52).

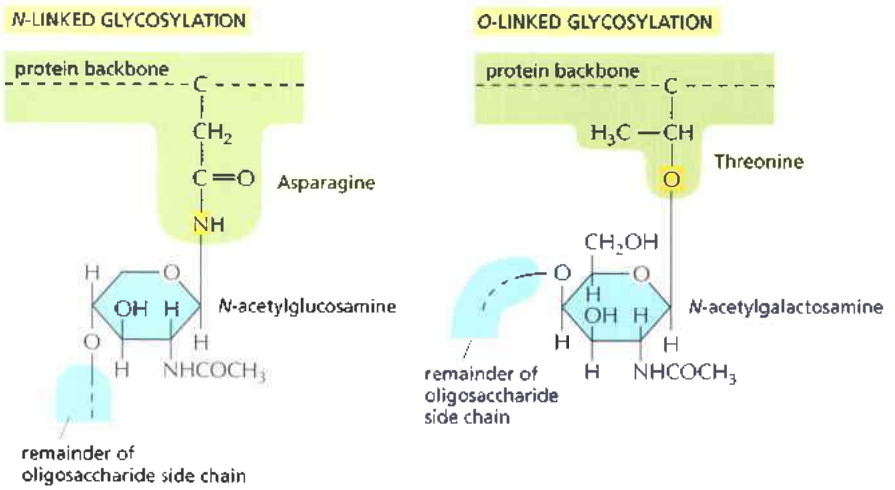


Figure 13–32 ***N*- and *O*-linked glycosylation.** In each case, only the single sugar group that is directly attached to the protein chain is shown.

add sugar residues to a protein one at a time. Usually, *N*-acetylgalactosamine is added first, followed by a variable number of additional sugar residues, ranging from just a few to 10 or more.

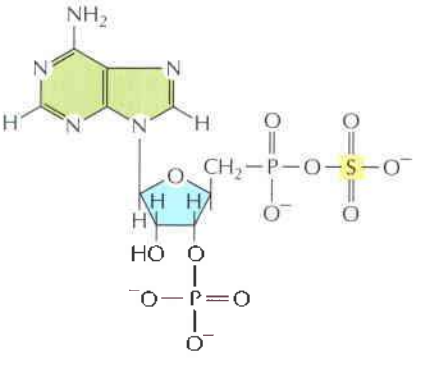
The Golgi apparatus confers the heaviest *O*-linked glycosylation of all on *mucins*, the glycoproteins in mucus secretions (see Figure 13–29), and on *proteoglycan core proteins*, which it modifies to produce **proteoglycans.** As discussed in Chapter 19, this process involves the polymerization of one or more *glycosaminoglycan chains* (long unbranched polymers composed of repeating disaccharide units; see Figure 19–58) via a xylose link onto serines on a core protein. Many proteoglycans are secreted and become components of the extracellular matrix, while others remain anchored to the extracellular face of the plasma membrane. Still others form a major component of slimy materials, such as the mucus that is secreted to form a protective coating on the surface of many epithelia.

The sugars incorporated into glycosaminoglycans are heavily sulfated in the Golgi apparatus immediately after these polymers are made, thus adding a significant portion of their characteristically large negative charge. Some tyrosines in proteins also become sulfated shortly before they exit from the Golgi apparatus. In both cases, the sulfation depends on the sulfate donor 3′-phosphoadenosine-5′-phosphosulfate (PAPS) (**Figure 13–33**), which is transported from the cytosol into the lumen of the *trans* Golgi network.

What Is the Purpose of Glycosylation?

There is an important difference between the construction of an oligosaccharide and the synthesis of other macromolecules such as DNA, RNA, and protein. Whereas nucleic acids and proteins are copied from a template in a repeated series of identical steps using the same enzyme or set of enzymes, complex carbohydrates require a different enzyme at each step, each product being recognized as the exclusive substrate for the next enzyme in the series. The vast abundance of glycoproteins and the complicated pathways that have evolved to synthesize them suggest that the oligosaccharides on glycoproteins and glycosphingolipids have very important functions.

N-linked glycosylation, for example, is prevalent in all eucaryotes, including yeasts. *N*-linked oligosaccharides also occur in a very similar form in archaeal cell wall proteins, suggesting that the whole machinery required for their synthesis is evolutionarily ancient. *N*-linked glycosylation promotes protein folding in two ways. First, it has a direct role in making folding intermediates more soluble, thereby preventing their aggregation. Second, the sequential modifications of the *N*-linked oligosaccharide establish a "glyco-code" that marks the progression of protein folding and mediates the binding of the protein to chaperones (discussed in Chapter 12) and lectins—for example, in guiding ER-to-Golgi transport. As we discuss later, lectins also participate in protein sorting in the *trans* Golgi network.



3′-phosphoadenosine-5′-phosphosulfate (PAPS)

Figure 13–33 **The structure of PAPS.**

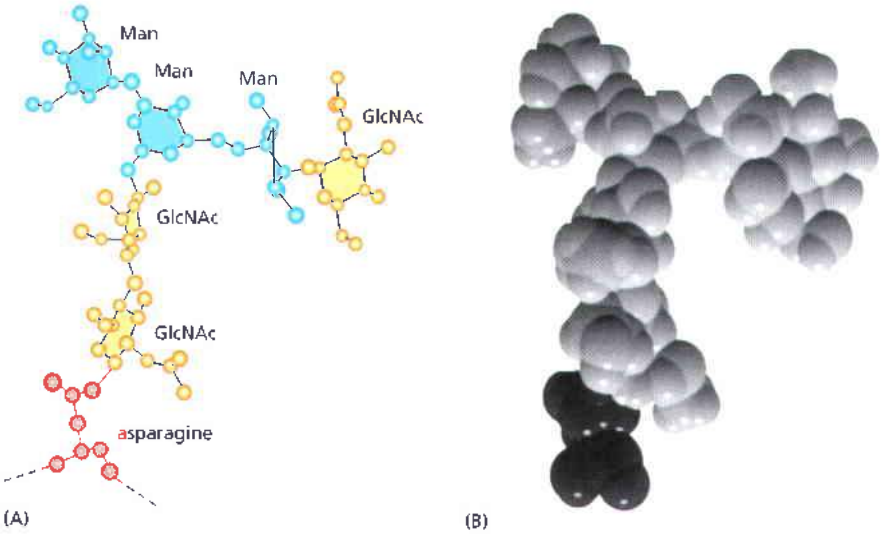


Figure 13–34 The three-dimensional structure of a small *N*-linked oligosaccharide. The structure was determined by x-ray crystallographic analysis of a glycoprotein. This oligosaccharide contains only 6 sugars, whereas there are 14 sugars in the *N*-linked oligosaccharide that is initially transferred to proteins in the ER (see Figures 12–50 and 12–51). (A) A backbone model showing all atoms except hydrogens. (B) A space-filling model, with the asparagine indicated by dark atoms. (B, courtesy of Richard Feldmann.)

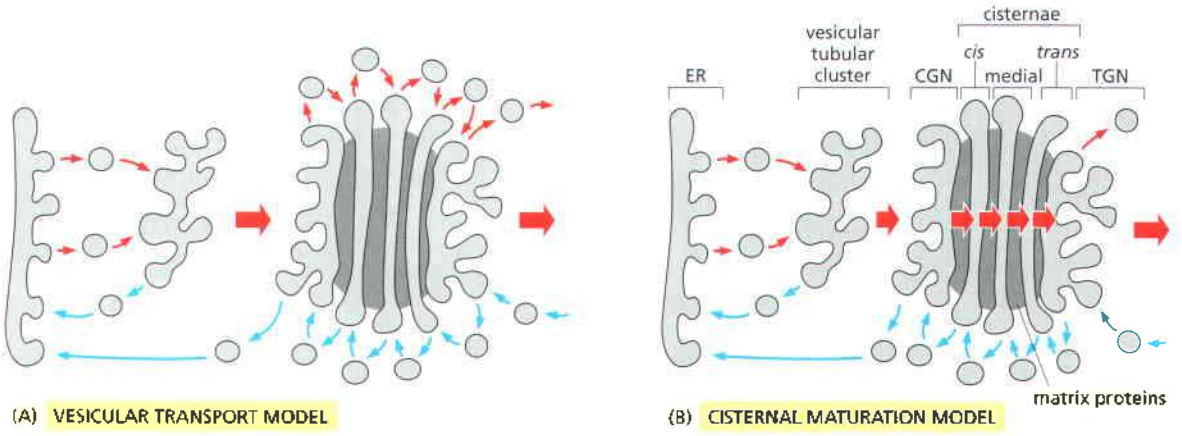
Because chains of sugars have limited flexibility, even a small *N*-linked oligosaccharide protruding from the surface of a glycoprotein (**Figure 13–34**) can limit the approach of other macromolecules to the protein surface. In this way, for example, the presence of oligosaccharides tends to make a glycoprotein more resistant to digestion by proteolytic enzymes. It may be that the oligosaccharides on cell-surface proteins originally provided an ancestral cell with a protective coat. Compared to the rigid bacterial cell wall, a mucus coat has the advantage that it leaves the cell with the freedom to change shape and move.

The sugar chains have since become modified to serve other purposes as well. The mucus coat of lung and intestinal cells, for example, protects against many pathogens. The recognition of sugar chains by *lectins* in the extracellular space is important in many developmental processes and in cell–cell recognition: *selectins*, for example, are lectins that function in cell–cell adhesion during lymphocyte migration, as discussed in Chapter 19. The presence of oligosaccharides may modify a protein's antigenic properties, making glycosylation an important factor in the production of proteins for pharmaceutical purposes.

Glycosylation can also have important regulatory roles. Signaling through the cell-surface signaling receptor Notch, for example, determines the cell's fate in development. Notch is a transmembrane protein that is *O*-glycosylated by addition of a single fucose to some serines, threonines, and hydroxylysines. Some cell types express an additional glycosyl transferase that adds an *N*-acetylglucosamine to each of these fucoses in the Golgi apparatus. This addition changes the specificity of the Notch receptor for the cell-surface signal proteins that activate the receptor.

Transport Through the Golgi Apparatus May Occur by Vesicular Transport or by Cisternal Maturation

It is still uncertain how the Golgi apparatus achieves and maintains its polarized structure and how molecules move from one cisterna to another. Functional evidence from *in vitro* transport assays, and the finding of abundant transport vesicles in the vicinity of Golgi cisternae, initially led to the view that these vesicles transport proteins between the cisternae, budding from one cisterna and fusing with the next. According to this **vesicular transport model**, the Golgi apparatus is a relatively static structure, with its enzymes held in place, while the molecules in transit move through the cisternae in sequence, carried by transport vesicles (**Figure 13–35**A). A retrograde flow of vesicles retrieves escaped ER and Golgi proteins and returns them to preceding compartments. Directional flow could be achieved because forward-moving cargo molecules are selectively allowed access to forward-moving vesicles. Although both forward-moving and retrograde types of vesicles are likely to be COPI-coated, the coats may contain different adaptor proteins that confer selectivity on the packaging of cargo



molecules. An alternative possibility is that the transport vesicles shuttling between Golgi cisternae are not directional at all, instead transporting cargo material randomly back and forth; directional flow would then occur because of the continual input at the *cis* cisterna and output at the *trans* cisterna.

A different hypothesis, called the **cisternal maturation model**, views the Golgi as a dynamic structure in which the cisternae themselves move. The vesicular tubular clusters that arrive from the ER fuse with one another to become a *cis* Golgi network. According to this model, this network then progressively matures to become a *cis* cisterna, then a medial cisterna, and so on. Thus, at the *cis* face of a Golgi stack, new *cis* cisternae would continually form and then migrate through the stack as they mature (Figure 13–35B). This model is supported by microscopic observations demonstrating that large structures such as collagen rods in fibroblasts and scales in certain algae—which are much too large to fit into classical transport vesicles—move progressively through the Golgi stack.

In the cisternal maturation model, retrograde flow explains the characteristic distribution of Golgi enzymes. Everything moves continuously forward with the maturing cisterna, including the processing enzymes that belong in the early Golgi apparatus. But budding COPI-coated vesicles continually collect the appropriate enzymes, almost all of which are membrane proteins, and carry them back to the earlier cisterna where they function. A newly formed *cis* cisterna would therefore receive its normal complement of resident enzymes primarily from the cisterna just ahead of it and would later pass those enzymes back to the next *cis* cisterna that forms.

As we discuss later, when a cisterna finally moves forward to become part of the *trans* Golgi network, various types of coated vesicles bud off it until this network disappears, to be replaced by a maturing cisterna just behind. At the same time, other transport vesicles are continually retrieving membrane from post-Golgi compartments and returning this membrane to the *trans* Golgi network.

The vesicular transport and the cisternal maturation models are not mutually exclusive. Indeed, evidence suggests that transport may occur by a combination of the two mechanisms, in which some cargo is moved forward rapidly in transport vesicles, whereas other cargo is moved forward more slowly as the Golgi apparatus constantly renews itself through cisternal maturation.

Figure 13–35 Two possible models explaining the organization of the Golgi apparatus and the transport of proteins from one cisterna to the next. It is likely that the transport through the Golgi apparatus in the forward direction *(red arrows)* involves elements of both models. (A) In the vesicular transport model, Golgi cisternae are static organelles, which contain a characteristic complement of resident enzymes. The passing of molecules from *cis* to *trans* through the Golgi is accomplished by forward-moving transport vesicles, which bud from one cisterna and fuse with the next in a *cis*-to-*trans* direction. (B) According to the alternative cisternal maturation model, each Golgi cisterna matures as it migrates outward through the stack. At each stage, the Golgi resident proteins that are carried forward in a cisterna are moved backward to an earlier compartment in COPI-coated vesicles. When a newly formed cisterna moves to a medial position, for example, "left-over" *cis* Golgi enzymes would be extracted and transported retrogradely to a new *cis* cisterna behind. Likewise, the medial enzymes would be received by retrograde transport from the cisternae just ahead. In this way, a *cis* cisterna would mature to a medial cisterna as it moves outward.

Golgi Matrix Proteins Help Organize the Stack

The unique architecture of the Golgi apparatus depends on both the microtubule cytoskeleton, as already mentioned, and cytoplasmic Golgi matrix proteins, which form a scaffold between adjacent cisternae and give the Golgi stack its structural integrity. Some of the matrix proteins form long, filamentous tethers, which are thought to help retain Golgi transport vesicles close to the organelle. When the cell prepares to divide, mitotic protein kinases phosphorylate the Golgi matrix proteins, causing the Golgi apparatus to fragment and disperse throughout the cytosol. The Golgi fragments are then distributed evenly to

the two daughter cells, where the matrix proteins are dephosphorylated, leading to the reassembly of the Golgi apparatus.

Remarkably, the Golgi matrix proteins can assemble into appropriately localized stacks near the centrosome even when Golgi membrane proteins are experimentally prevented from leaving the ER. This observation suggests that the matrix proteins are largely responsible for both the structure and location of the Golgi apparatus.

Summary

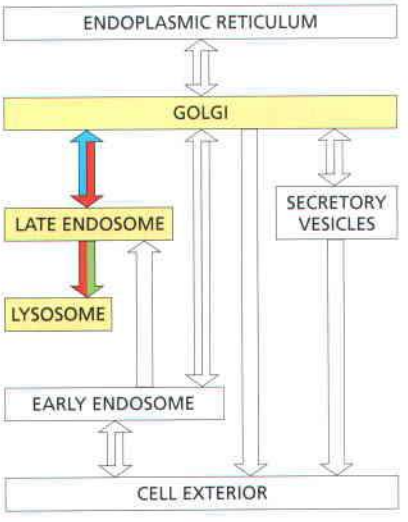
Correctly folded and assembled proteins in the ER are packaged into COPII-coated transport vesicles that pinch off from the ER membrane. Shortly thereafter, the vesicles shed their coat and fuse with one another to form vesicular tubular clusters. The clusters then move on microtubule tracks to the Golgi apparatus, where they fuse with one another to form the cis*-Golgi network. Any resident ER proteins that escape from the ER are returned there from the vesicular tubular clusters and Golgi apparatus by retrograde transport in COPI-coated vesicles.*

The Golgi apparatus, unlike the ER, contains many sugar nucleotides, which glycosyl transferase enzymes use to perform glycosylation reactions on lipid and protein molecules as they pass through the Golgi apparatus. The mannoses on the N-linked oligosaccharides that are added to proteins in the ER are often initially removed, and further sugars are added. Moreover, the Golgi apparatus is the site where O-linked glycosylation occurs and where glycosaminoglycan chains are added to core proteins to form proteoglycans. Sulfation of the sugars in proteoglycans and of selected tyrosines on proteins also occurs in a late Golgi compartment.

The Golgi apparatus modifies the many proteins and lipids that it receives from the ER and then distributes them to the plasma membrane, lysosomes, and secretory vesicles. The Golgi apparatus is a polarized organelle, consisting of one or more stacks of disc-shaped cisternae, each stack organized as a series of at least three functionally distinct compartments, termed cis, medial, *and* trans *cisternae. The* cis *and* trans *cisternae are both connected to special sorting stations, called the* cis *Golgi network and the* trans *Golgi network, respectively. Proteins and lipids move through the Golgi stack in the* cis*-to-*trans *direction. This movement may occur by vesicular transport, by progressive maturation of the* cis *cisternae as they migrate continuously through the stack, or, most likely, by a combination of these two mechanisms. Continual retrograde vesicular transport from more distal cisternae is thought to keep the enzymes concentrated in the cisternae where they are needed. The finished new proteins end up in the* trans *Golgi network, which packages them in transport vesicles and dispatches them to their specific destinations in the cell.*

TRANSPORT FROM THE *TRANS* GOLGI NETWORK TO LYSOSOMES

The *trans* Golgi network sorts all of the proteins that pass through the Golgi apparatus (except those that are retained there as permanent residents) according to their final destination. The sorting mechanism is especially well understood for those proteins destined for the lumen of lysosomes, and in this section we consider this selective transport process. We begin with a brief account of lysosome structure and function.



Lysosomes Are the Principal Sites of Intracellular Digestion

Lysosomes are membrane-enclosed compartments filled with soluble hydrolytic enzymes that control intracellular digestion of macromolecules. Lysosomes contain about 40 types of hydrolytic enzymes, including proteases, nucleases, glycosidases, lipases, phospholipases, phosphatases, and sulfatases. All are **acid hydrolases**. For optimal activity, they need to be activated by

proteolytic cleavage and require an acid environment, which the lysosome provides by maintaining a pH of about 4.5–5.0 in its interior. By this arrangement, the contents of the cytosol are doubly protected against attack by the cell's own digestive system: the membrane of the lysosome keeps the digestive enzymes out of the cytosol, but even if they leak out, they can do little damage at the cytosolic pH of about 7.2.

Like all other intracellular organelles, the lysosome not only contains a unique collection of enzymes, but also has a unique surrounding membrane. Most of the lysosomal membrane proteins, for example, are unusually highly glycosylated, which helps to protect them from the lysosomal proteases in the lumen. Transport proteins in the lysosomal membrane carry the final products of the digestion of macromolecules—such as amino acids, sugars, and nucleotides—to the cytosol, where the cell can either reuse or excrete them.

A *vacuolar H⁺ ATPase* in the lysosomal membrane uses the energy of ATP hydrolysis to pump H^+ into the lysosome, thereby maintaining the lumen at its acidic pH (Figure 13–36). The lysosomal H^+ pump belongs to the family of *V-type ATPases* and has a similar architecture to the mitochondrial and chloroplast ATP synthases (F-type ATPases), which convert the energy stored in H^+ gradients into ATP (see Figure 11–12). By contrast to these enzymes, however, the vacuolar H^+ ATPase exclusively works in reverse, pumping H^+ into the organelle. Similar or identical V-type ATPases acidify all endocytic and exocytic organelles, including lysosomes, endosomes, selected compartments of the Golgi apparatus, and many transport and secretory vesicles. In addition to providing a low-pH environment that is suitable for reactions occurring in the organelle lumen, the H^+ gradient provides a source of energy that drives the transport of small metabolites across the organelle membrane.

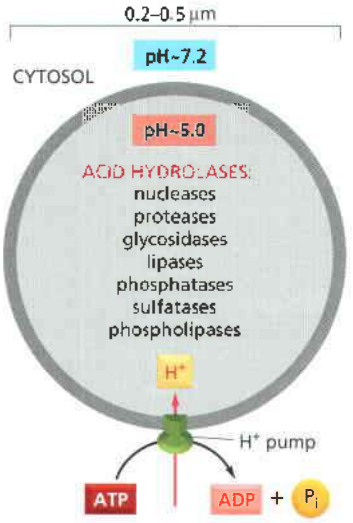


Figure 13–36 **Lysosomes.** The acid hydrolases are hydrolytic enzymes that are active under acidic conditions. A V-type ATPase in the membrane pumps H^+ into the lysosome, maintaining its lumen at an acidic pH.

Lysosomes Are Heterogeneous

Lysosomes were initially discovered by the biochemical fractionation of cell extracts; only later were they seen clearly in the electron microscope. Although extraordinarily diverse in shape and size, staining them with specific antibodies shows they are members of a single family of organelles. They can also be identified by histochemistry, using the precipitate produced by the action of an acid hydrolase on its substrate to indicate which organelles contain the hydrolase (Figure 13–37). By this criterion, lysosomes are found in all eucaryotic cells.

The heterogeneity of lysosomal morphology contrasts with the relatively uniform structures of most other cell organelles. The diversity reflects the wide variety of digestive functions that acid hydrolases mediate, including the breakdown of intra- and extracellular debris, the destruction of phagocytosed microorganisms, and the production of nutrients for the cell. The diversity of lysosomal morphology, however, also reflects the way lysosomes form: late endosomes contain material received from both the plasma membrane by endocytosis and newly synthesized lysosomal hydrolases, and they therefore already bear a resemblance to lysosomes. Late endosomes fuse with preexisting lysosomes to form structures that are sometimes referred to as *endolysosomes*, which then fuse with one another (Figure 13–38). When the majority of the endocytosed material within an endolysosome has been digested so that only resistant or slowly digestable residues remain, these organelles become "classical" lysosomes. These are relatively dense, round, and small, but they can enter

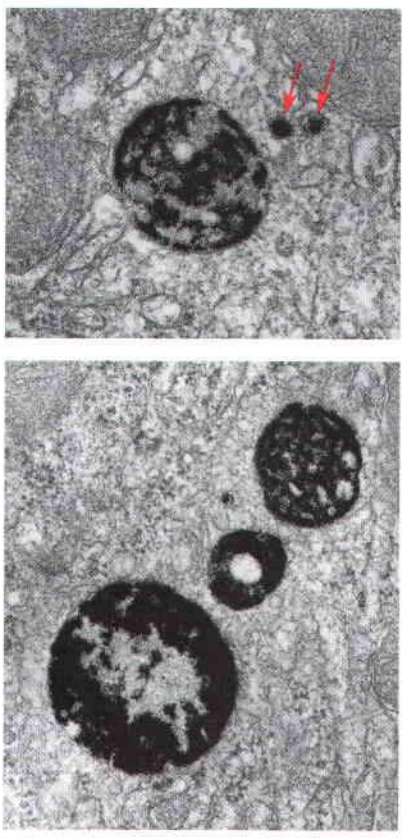


Figure 13–37 **Histochemical visualization of lysosomes.** These electron micrographs show two sections of a cell stained to reveal the location of acid phosphatase, a marker enzyme for lysosomes. The larger membrane-enclosed organelles, containing dense precipitates of lead phosphate, are lysosomes. Their diverse morphology reflects variations in the amount and nature of the material they are digesting. The precipitates are produced when tissue fixed with glutaraldehyde (to fix the enzyme in place) is incubated with a phosphatase substrate in the presence of lead ions. *Red arrows* in the top panel indicate two small vesicles thought to be carrying acid hydrolases from the Golgi apparatus. (Courtesy of Daniel S. Friend.)

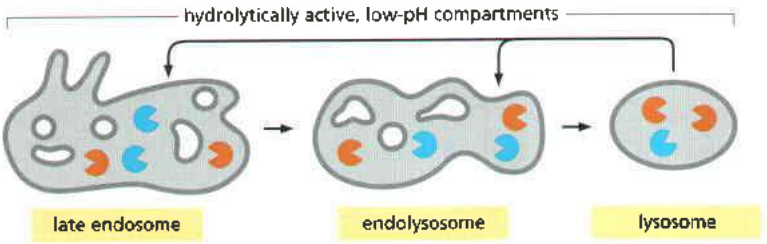


Figure 13–38 **A model for lysosome maturation.** The heterogeneity of lysosomal morphology reflects, in part, the different nature of the materials delivered to the organelle, as well as the different stages in the maturation cycle shown here.

the cycle again by fusing with late endosomes or endolysosomes. Thus, there is no real distinction between late endosomes and lysosomes: they are the same except that they are in different stages of a maturation cycle. For this reason, lysosomes are sometimes viewed as a heterogeneous collection of distinct organelles, the common feature of which is a high content of hydrolytic enzymes. It is especially hard to apply a narrower definition than this in plant cells, as we discuss next.

Plant and Fungal Vacuoles Are Remarkably Versatile Lysosomes

Most plant and fungal cells (including yeasts) contain one or several very large, fluid-filled vesicles called **vacuoles**. They typically occupy more than 30% of the cell volume, and as much as 90% in some cell types (Figure 13–39). Vacuoles are related to animal cell lysosomes and contain a variety of hydrolytic enzymes, but their functions are remarkably diverse. The plant vacuole can act as a storage organelle for both nutrients and waste products, as a degradative compartment, as an economical way of increasing cell size (Figure 13–40), and as a controller of *turgor pressure* (the osmotic pressure that pushes outward on the cell wall and keeps the plant from wilting). The same cell may have different vacuoles with distinct functions, such as digestion and storage.

The vacuole is important as a homeostatic device, enabling plant cells to withstand wide variations in their environment. When the pH in the environment drops, for example, the flux of H^+ into the cytosol is balanced, at least in part, by an increased transport of H^+ into the vacuole, which tends to keep the pH in the cytosol constant. Similarly, many plant cells maintain an almost constant turgor pressure despite large changes in the tonicity of the fluid in their

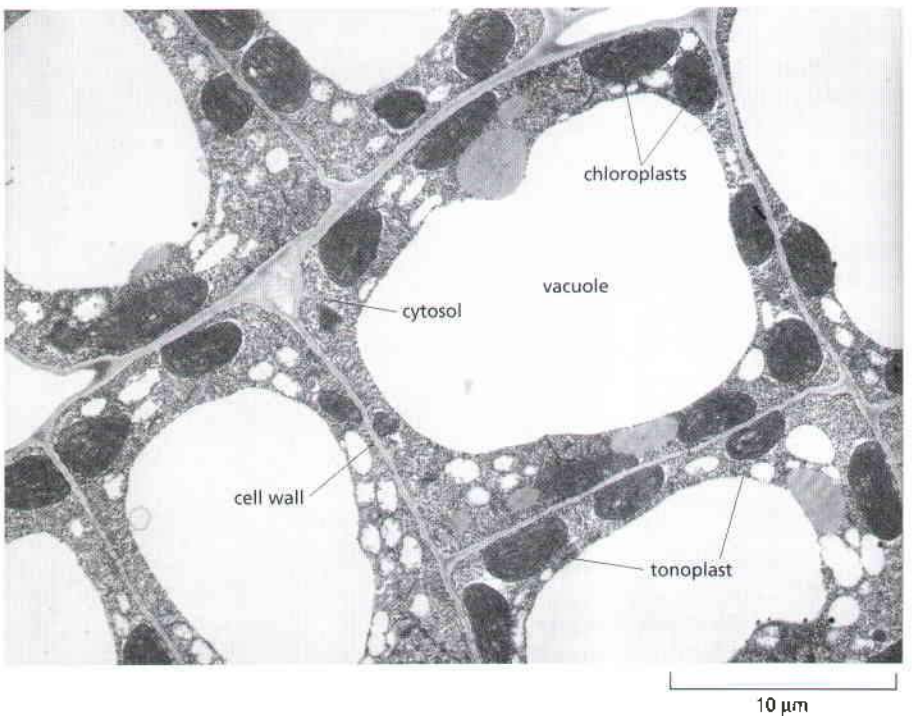


Figure 13–39 **The plant cell vacuole.** This electron micrograph of cells in a young tobacco leaf shows the cytosol as a thin layer, containing chloroplasts, pressed against the cell wall by the enormous vacuole. The vacuole membrane is called the tonoplast. (Courtesy of J. Burgess.)

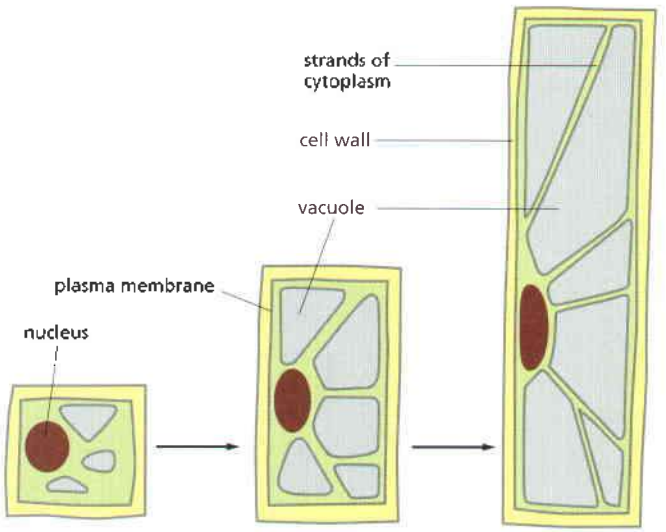


Figure 13–40 **The role of the vacuole in controlling the size of plant cells.** A plant cell can achieve a large increase in volume without increasing the volume of the cytosol. Localized weakening of the cell wall orients a turgor-driven cell enlargement that accompanies the uptake of water into an expanding vacuole. The cytosol is eventually confined to a thin peripheral layer, which is connected to the nuclear region by strands of cytoplasm stabilized by bundles of actin filaments (not shown).

immediate environment. They do so by changing the osmotic pressure of the cytosol and vacuole—in part by the controlled breakdown and resynthesis of polymers such as polyphosphate in the vacuole, and in part by altering the transport rates of sugars, amino acids, and other metabolites across the plasma membrane and the vacuolar membrane. The turgor pressure regulates the activities of distinct transporters in each membrane to control these fluxes.

Humans often harvest substances stored in plant vacuoles. In different species, these range from rubber to opium to the flavoring of garlic. Many stored products have a metabolic function. Proteins, for example, can be preserved for years in the vacuoles of the storage cells of many seeds, such as those of peas and beans. When the seeds germinate, these proteins are hydrolyzed and the resulting amino acids provide a food supply for the developing embryo. Anthocyanin pigments stored in vacuoles color the petals of many flowers so as to attract pollinating insects, while noxious molecules released from vacuoles when a plant is eaten or damaged provide a defense against predators.

Multiple Pathways Deliver Materials to Lysosomes

Lysosomes are usually meeting places where several streams of intracellular traffic converge. A route that leads outwards from the ER via the Golgi apparatus delivers most digestive enzymes, while at least three paths from different sources feed substances into lysosomes for digestion.

The best studied of these degradation paths in lysosomes is the one followed by macromolecules taken up from extracellular fluid by *endocytosis*. As discussed in detail later, endocytosed molecules are initially delivered in vesicles to small, irregularly shaped intracellular organelles called *early endosomes*. Here endocytosed materials first meet the lysosomal hydrolases, which are delivered to the endosome from the Golgi apparatus. Some of the ingested molecules are selectively retrieved and recycled to the plasma membrane, while others pass on into *late endosomes*. The interior of the late endosomes is mildly acidic (pH ~6), and it is the site where the hydrolytic digestion of the endocytosed molecules begins. As discussed above, mature lysosomes form by a maturation process from late endosomes, accompanied by a further decrease in internal pH. As lysosomes mature, endosomal membrane proteins are selectively retrieved from the developing lysosome by transport vesicles that deliver these proteins back to endosomes or the TGN.

All cell types use a second degradation pathway in lysosomes to dispose of obsolete parts of the cell itself—a process called **autophagy**. In a liver cell, for example, an average mitochondrion has a lifetime of about 10 days, and electron microscopic images of normal cells reveal lysosomes containing (and presumably digesting) mitochondria, as well as other organelles. The process seems to begin with the enclosure of an organelle by a double membrane of

unknown origin, creating an *autophagosome*, which then fuses with a lysosome (or a late endosome). The process is highly regulated, and selected cell components can somehow be marked for lysosomal destruction during cell remodeling. For example, the smooth ER that proliferates in a liver cell in a detoxification response to lipid-soluble drugs such as phenobarbital (discussed in Chapter 12) is selectively removed by autophagy after the drug is withdrawn.

Similarly, other obsolete organelles, including senescent peroxisomes or mitochondria, can be selectively targeted for degradation by autophagy. Under starvation conditions, large portions of the cytosol are nonselectively captured into autophagosomes. Metabolites derived from the digestion of the captured material help the cell survive when external nutrients are limiting.

In addition to maintaining basic cell functions in balance and helping to dispose of obsolete parts, autophagy also has a role in development and health. It helps restructure differentiating cells by disposing of no longer needed parts and helps defend against invading viruses and bacteria. Autophagy is uniquely suited as a mechanism that can remove whole organelles or large protein aggregates, which other mechanisms such as proteasomal degradation cannot handle.

We still know very little about the events that lead to the formation of autophagosomes, or how the process of autophagy is controlled and targeted at specific organelles. More than 25 different proteins have been identified in yeast and animal cells that participate in the process. Autophagy can be divided into four general steps: (1) nucleation and extension of a delimiting membrane into a crescent-shaped structure that engulfs a portion of the cytoplasm, (2) closure of the autophagosome into a sealed double-membrane-bounded compartment, (3) fusion of the new compartment with lysosomes, and (4) digestion of the inner membrane of the autophagosome and its contents (**Figure 13–41**). Many mysteries remain to be solved, including identifying the membrane system from which the vesicles that form the autophagosomal envelope derive, and how some target organelles can be enclosed so selectively.

As we discuss later, the third pathway that brings materials to lysosomes for degradation is found mainly in cells specialized for the *phagocytosis* of large particles and microorganisms. Such professional phagocytes (macrophages and neutrophils in vertebrates) engulf objects to form a *phagosome*, which is then converted to a lysosome in the manner described for the autophagosome. **Figure 13–42** summarizes the three pathways.

A Mannose 6-Phosphate Receptor Recognizes Lysosomal Proteins in the *Trans* Golgi Network

We now consider the pathway that delivers lysosomal hydrolases and membrane proteins to lysosomes. Both classes of proteins are co-translationally transported into the rough ER and then transported through the Golgi apparatus to the TGN. The transport vesicles that deliver these proteins to endosomes (from where the proteins are moved on to lysosomes) bud from the TGN. The vesicles incorporate the lysosomal proteins and exclude the many other proteins being packaged into different transport vesicles for delivery elsewhere.

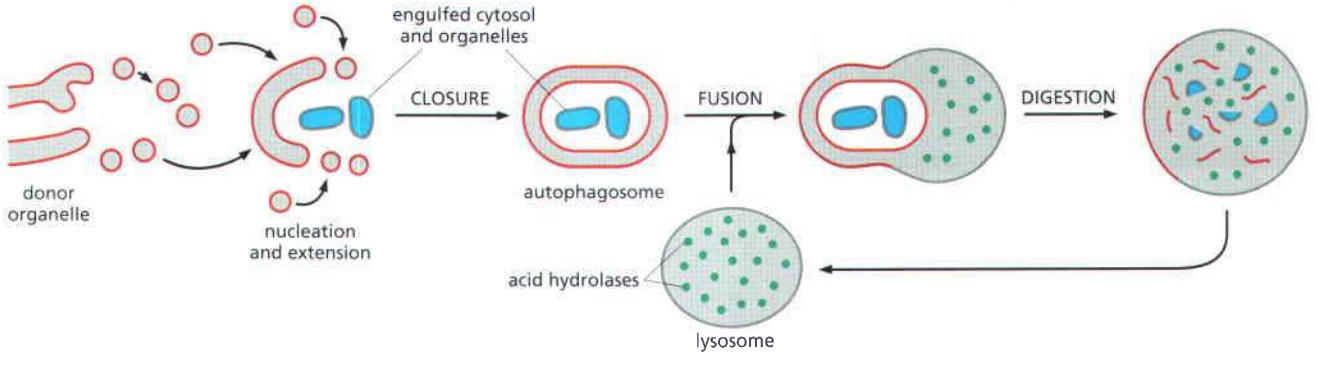
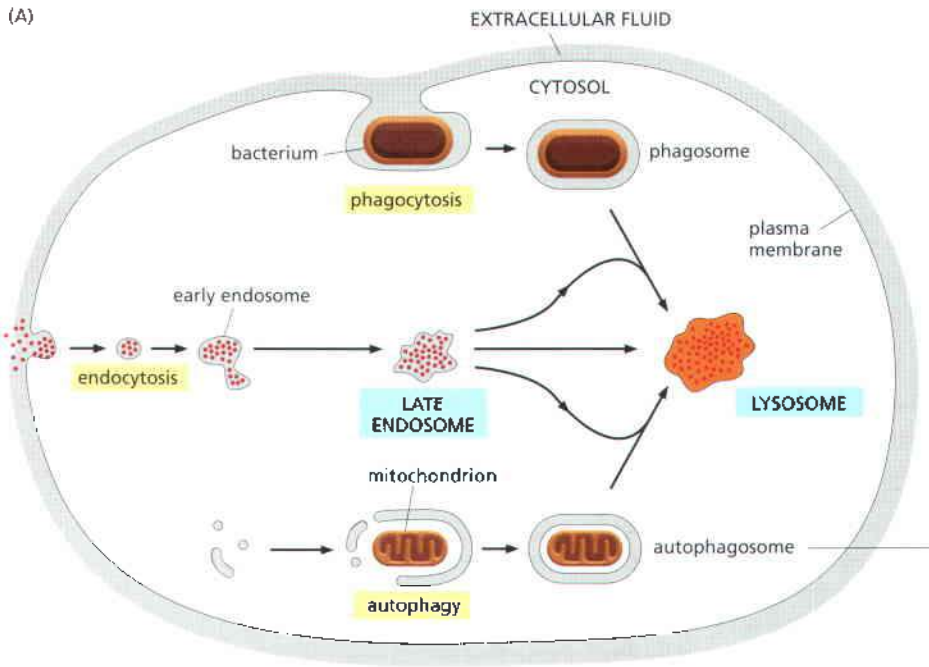


Figure 13–41 **A model of autophagy.** After a nucleation event in the cytoplasm, a crescent of autophagosomal membrane grows by fusion of vesicles of unknown origin that extend its edges. Eventually, a membrane fusion event closes the autophagosome, sequestering a portion of the cytoplasm of the cell in a double membrane. The autophagosome then fuses with lysosomes containing acid hydrolases that digest its content.



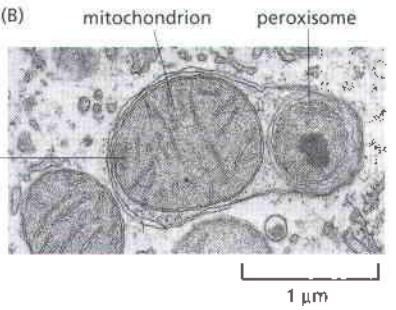


Figure 13–42 **Three pathways to degradation in lysosomes.** (A) Materials in each pathway are derived from a different source. Note that the autophagosome has a double membrane. (B) An electron micrograph of an autophagosome containing a mitochondrion and a peroxisome. (B, courtesy of Daniel S. Friend, from D.W. Fawcett, A Textbook of Histology, 12th ed. New York: Chapman and Hall, 1994. With permission from Kluwer.)

How are lysosomal proteins recognized and selected in the TGN with the required accuracy? We know the answer for the lysosomal hydrolases. They carry a unique marker in the form of *mannose 6-phosphate (M6P)* groups, which are added exclusively to the *N*-linked oligosaccharides of these soluble lysosomal enzymes as they pass through the lumen of the *cis* Golgi network (**Figure 13–43**). Transmembrane **M6P receptor proteins**, which are present in the TGN, recognize the M6P groups. The receptor proteins bind to lysosomal hydrolases on the lumenal side of the membrane and to adaptor proteins in assembling clathrin coats on the cytosolic side. In this way, they help package the hydrolases into clathrin-coated vesicles that bud from the TGN. The vesicles shed their coat and deliver their contents to early endosomes.

The M6P Receptor Shuttles Between Specific Membranes

The M6P receptor protein binds its specific oligosaccharide at pH 6.5–6.7 in the TGN and releases it at pH 6, which is the pH in the interior of late endosomes. Thus, as the pH drops during endosomal maturation, the lysosomal hydrolases dissociate from the M6P receptor and eventually begin to digest the material delivered by endocytosis. An acid phosphatase removes the phosphate group from the mannose, thereby destroying the sorting signal and contributing to the release of the lysosomal hydrolases from the M6P receptor. Having released their bound enzymes, the M6P receptors are retrieved into retromer-coated transport vesicles that bud from endosomes; the receptors are then returned to the membrane of the TGN for reuse (**Figure 13–44**). Transport in either direction requires signals in the cytoplasmic tail of the M6P receptor that direct this protein to the endosome or back to the Golgi apparatus. These signals are recognized by the retromer complex (see Figure 13–9) that recruits M6P receptors into vesicles in endosomes. The recycling of the M6P receptor resembles the recycling of the KDEL receptor discussed earlier, although it differs in the type of coated vesicles that mediate the transport.

Not all the hydrolase molecules that are tagged with M6P get to lysosomes. Some escape the normal packaging process in the *trans* Golgi network and are transported "by default" to the cell surface, where they are secreted into the extracellular fluid. Some M6P receptors, however, also take a detour to the plasma membrane, where they recapture the escaped lysosomal hydrolases and return them by *receptor-mediated endocytosis* to lysosomes via early and late endosomes. As lysosomal hydrolases require an acidic milieu to work, they can do little harm in the extracellular fluid, which usually has a neutral pH of 7.4.

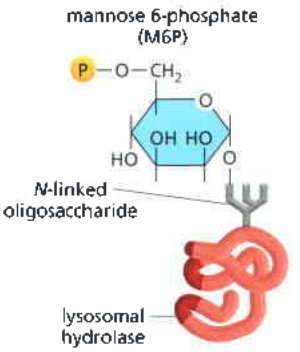
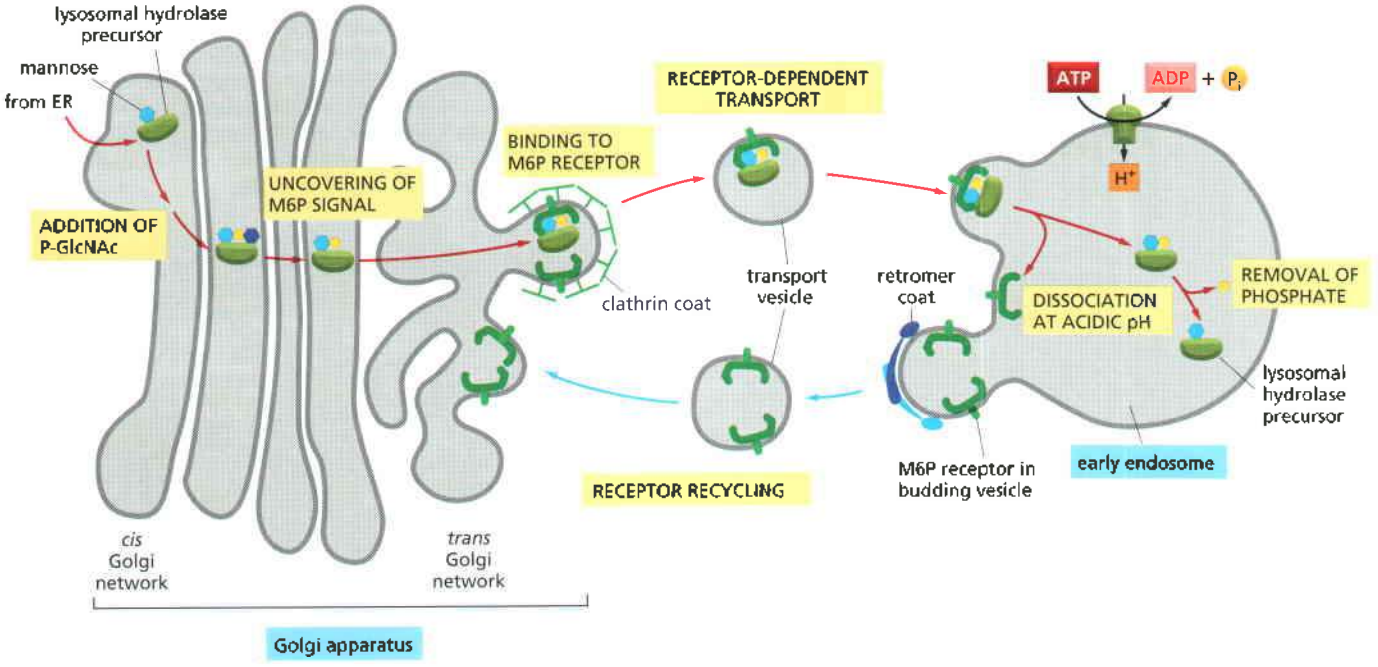


Figure 13–43 **The structure of mannose 6-phosphate on a lysosomal hydrolase.**



A Signal Patch in the Hydrolase Polypeptide Chain Provides the Cue for M6P Addition

The sorting system that segregates lysosomal hydrolases and dispatches them to endosomes works because M6P groups are added only to the appropriate glycoproteins in the Golgi apparatus. This requires specific recognition of the hydrolases by the Golgi enzymes responsible for adding M6P. Since all glycoproteins leave the ER with identical *N*-linked oligosaccharide chains, the signal for adding the M6P units to oligosaccharides must reside somewhere in the polypeptide chain of each hydrolase. Genetic engineering experiments have revealed that the recognition signal is a cluster of neighboring amino acids on each protein's surface, known as a *signal patch.*

Two enzymes act sequentially to catalyze the addition of M6P groups to lysosomal hydrolases. The first is a GlcNAc phosphotransferase in the *cis* Golgi that specifically binds the hydrolase and adds GlcNAc-phosphate to one or two of the mannose residues on each oligosaccharide chain (Figure 13–45). A second enzyme in the *trans* Golgi then cleaves off the GlcNAc residue, leaving behind a newly created M6P marker. Since most lysosomal hydrolases contain multiple oligosaccharides, they acquire many M6P residues, providing a high-affinity signal for the M6P receptor.

Defects in the GlcNAc Phosphotransferase Cause a Lysosomal Storage Disease in Humans

Genetic defects that affect one or more of the lysosomal hydrolases cause a number of human **lysosomal storage diseases**. The defects result in an accumulation of undigested substrates in lysosomes, with severe pathological consequences, most often in the nervous system. In most cases, there is a mutation in a structural gene that codes for an individual lysosomal hydrolase. This occurs in *Hurler's disease*, for example, in which the enzyme required for the breakdown of certain types of glycosaminoglycan chains is defective or missing. The most severe form of lysosomal storage disease, however, is a very rare disorder called *inclusion-cell disease* (*I-cell disease*). In this condition, almost all of the hydrolytic enzymes are missing from the lysosomes of fibroblasts, and their undigested substrates accumulate in lysosomes, which consequently form large *inclusions* in the patients' cells.

Figure 13–44 The transport of newly synthesized lysosomal hydrolases to lysosomes. The sequential action of two enzymes in the *cis* and *trans* Golgi network adds mannose 6-phosphate (M6P) groups to the precursors of lysosomal enzymes (see Figure 13–45). They then segregate from all other types of proteins in the TGN because monomeric adaptor proteins in the clathrin coat bind the M6P receptors, which, in turn, bind the modified lysosomal hydrolases. The clathrin-coated vesicles bud off from the TGN, shed their coat, and fuse with early endosomes. At the lower pH of the endosome, the hydrolases dissociate from the M6P receptors, and the empty receptors are recycled in retromer-coated vesicles to the Golgi apparatus for further rounds of transport. In the endosomes, the phosphate is removed from the mannose sugars attached to the hydrolases, further ensuring that the hydrolases do not return to the Golgi apparatus with the receptor.

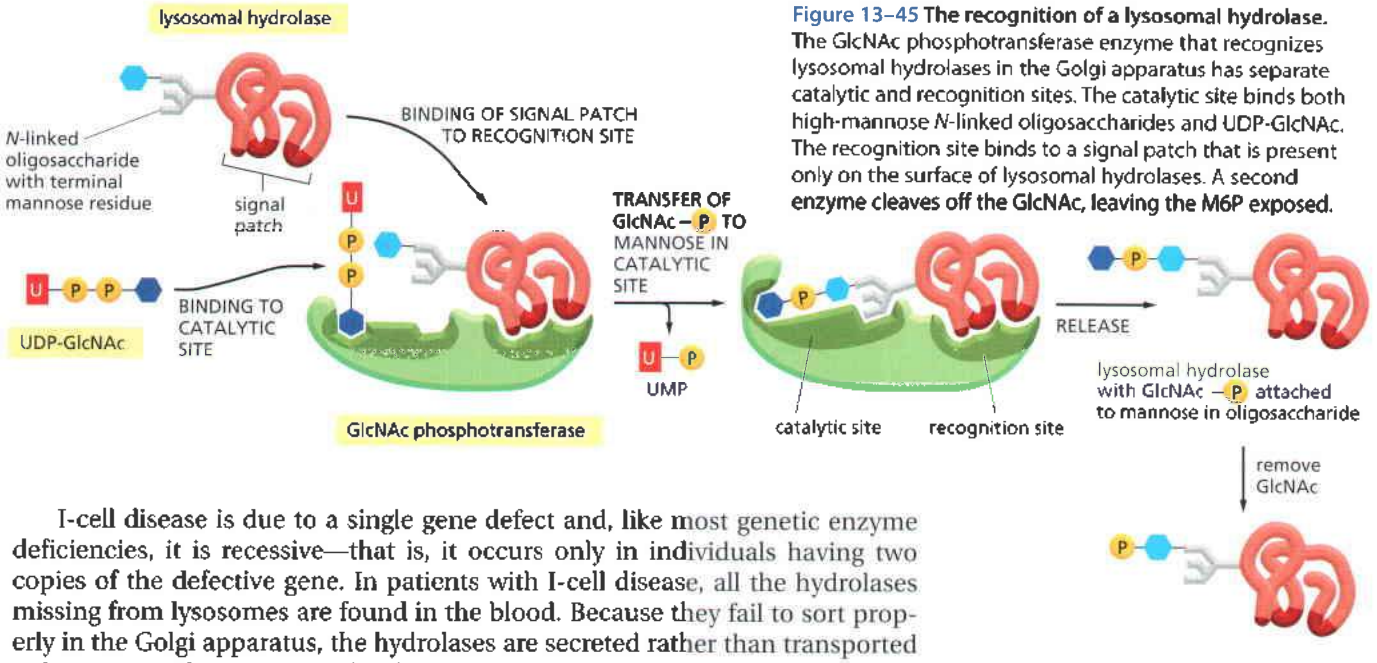


Figure 13–45 The recognition of a lysosomal hydrolase. The GlcNAc phosphotransferase enzyme that recognizes lysosomal hydrolases in the Golgi apparatus has separate catalytic and recognition sites. The catalytic site binds both high-mannose *N*-linked oligosaccharides and UDP-GlcNAc. The recognition site binds to a signal patch that is present only on the surface of lysosomal hydrolases. A second enzyme cleaves off the GlcNAc, leaving the M6P exposed.

I-cell disease is due to a single gene defect and, like most genetic enzyme deficiencies, it is recessive—that is, it occurs only in individuals having two copies of the defective gene. In patients with I-cell disease, all the hydrolases missing from lysosomes are found in the blood. Because they fail to sort properly in the Golgi apparatus, the hydrolases are secreted rather than transported to lysosomes. The missorting has been traced to a defective or missing GlcNAc-phosphotransferase. Because lysosomal enzymes are not phosphorylated in the *cis* Golgi network, the M6P receptors do not segregate them into the appropriate transport vesicles in the TGN. Instead, the lysosomal hydrolases are carried to the cell surface and secreted by a default pathway.

In I-cell disease, the lysosomes in some cell types, such as hepatocytes, contain a normal complement of lysosomal enzymes, implying that there is another pathway for directing hydrolases to lysosomes that is used by some cell types but not others. The nature of this M6P-independent pathway is unknown. Similarly, an M6P-independent pathway in all cells sorts the membrane proteins of lysosomes from the TGN for transport to late endosomes, and those proteins are therefore normal in I-cell disease. These membrane proteins exit from the TGN in clathrin-coated vesicles that are distinct from those that transport the M6P-tagged hydrolases and use different adaptor proteins.

It is unclear why cells need more than one sorting pathway to construct lysosomes, although it is perhaps not surprising that different mechanisms operate for soluble and membrane-bound lysosomal proteins, especially since—unlike M6P receptors—those membrane proteins are lysosomal residents and need not be returned to the TGN.

Some Lysosomes Undergo Exocytosis

Targeting of material to lysosomes is not necessarily the end of the pathway. *Lysosomal secretion* of undigested content enables all cells to eliminate undigestible debris. For most cells, this seems to be a minor pathway, used only when the cells are stressed. Some cell types, however, contain specialized lysosomes that have acquired the necessary machinery for fusion with the plasma membrane. *Melanocytes* in the skin, for example, produce and store pigments in their lysosomes. These pigment-containing *melanosomes* release their pigment into the extracellular space of the epidermis by exocytosis. The pigment is then taken up by keratinocytes, leading to normal skin pigmentation. In some genetic disorders, defects in melanosome exocytosis block this transfer process, leading to forms of hypopigmentation (albinism).

Summary

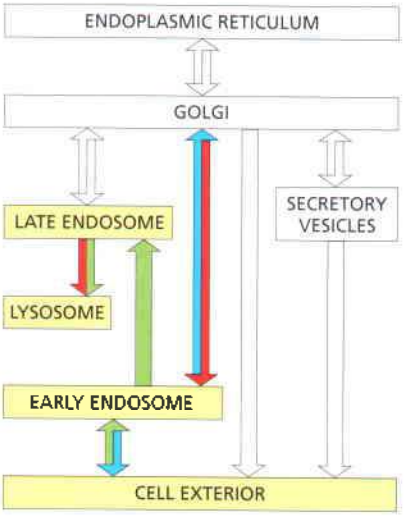
Lysosomes are specialized for the intracellular digestion of macromolecules. They contain unique membrane proteins and a wide variety of soluble hydrolytic enzymes that

operate best at pH 5, which is the internal pH of lysosomes. An ATP-driven H^+ pump in the lysosomal membrane maintains this low pH. Newly synthesized lysosomal proteins are transferred into the lumen of the ER, transported through the Golgi apparatus, and then carried from the trans *Golgi network to late endosomes by means of clathrin-coated transport vesicles.*

The lysosomal hydrolases contain N*-linked oligosaccharides that are covalently modified in a unique way in the* cis *Golgi network so that their mannose residues are phosphorylated. These mannose 6-phosphate (M6P) groups are recognized by an M6P receptor protein in the* trans *Golgi network that segregates the hydrolases and helps package them into budding transport vesicles that deliver their contents to endosomes. The M6P receptors shuttle back and forth between the* trans *Golgi network and these endosomes. The low pH in endosomes and removal of the phosphate from the M6P group cause the lysosomal hydrolases to dissociate from these receptors, making the transport of the hydrolases unidirectional. A separate transport system uses clathrin-coated vesicles to deliver resident lysosomal membrane proteins from the* trans *Golgi network.*

TRANSPORT INTO THE CELL FROM THE PLASMA MEMBRANE: ENDOCYTOSIS

The routes that lead inward from the cell surface to lysosomes start with the process of **endocytosis**, by which cells take up macromolecules, particulate substances, and, in specialized cases, even other cells. In this process, the material to be ingested is progressively enclosed by a small portion of the plasma membrane, which first invaginates and then pinches off to form an *endocytic vesicle* containing the ingested substance or particle. Two main types of endocytosis differ according to the size of the endocytic vesicles formed. In *phagocytosis* ("cell eating"), large particles are ingested via large vesicles called *phagosomes* (generally >250 nm in diameter). In *pinocytosis* ("cell drinking"), fluid and solutes are ingested via small *pinocytic vesicles* (about 100 nm in diameter). Most eucaryotic cells are continually ingesting fluid and solutes by pinocytosis; large particles are ingested most efficiently by specialized phagocytic cells.



Specialized Phagocytic Cells Can Ingest Large Particles

Phagocytosis is a special form of endocytosis in which a cell uses large endocytic vesicles called **phagosomes** to ingest large particles such as microorganisms and dead cells. <TCAT> In protozoa, phagocytosis is a form of feeding: large particles taken up into phagosomes end up in lysosomes, and the products of the subsequent digestive processes pass into the cytosol to be used as food. However, few cells in multicellular organisms are able to ingest such large particles efficiently. In the gut of animals, for example, extracellular processes break down food particles, and cells import the small hydrolysis products.

Phagocytosis is important in most animals for purposes other than nutrition, and it is carried out mainly by specialized cells—so-called *professional phagocytes*. In mammals, two classes of white blood cells act as professional phagocytes—**macrophages** and **neutrophils**. These cells develop from hemopoietic stem cells (discussed in Chapter 23), and they ingest invading microorganisms to defend us against infection. Macrophages also have an important role in scavenging senescent cells and cells that have died by apoptosis (discussed in Chapter 18). In quantitative terms, the clearance of senescent and dead cells is by far the most important: our macrophages phagocytose more than 10^{11} senescent red blood cells in each of us every day, for example.

Whereas the endocytic vesicles involved in pinocytosis are small and uniform, phagosomes have diameters that are determined by the size of the ingested particle, and they can be almost as large as the phagocytic cell itself (**Figure 13–46**). The phagosomes fuse with lysosomes inside the cell, and the ingested material is then degraded. Any indigestible substances will remain in

lysosomes, forming *residual bodies*, which can be excreted from cells by exocytosis, as we have previously discussed. Some of the internalized plasma membrane components never reach the lysosome, because they are retrieved from the phagosome in transport vesicles and returned to the plasma membrane.

To be phagocytosed, particles must first bind to the surface of the phagocyte. Not all particles that bind are ingested, however. Phagocytes have a variety of specialized surface receptors that are functionally linked to the phagocytic machinery of the cell. Phagocytosis is a triggered process. That is, it requires the activation of receptors that transmit signals to the cell interior and initiate the response. By contrast, pinocytosis is a constitutive process. It occurs continuously, regardless of the needs of the cell. The best-characterized triggers of phagocytosis are antibodies, which protect us by binding to the surface of infectious microorganisms to form a coat that exposes the tail region on the exterior of each antibody molecule. This tail region is called the Fc region (discussed in Chapter 25). The antibody coat is recognized by specific *Fc receptors* on the surface of macrophages and neutrophils, whose binding induces the phagocytic cell to extend pseudopods that engulf the particle and fuse at their tips to form a phagosome (**Figure 13–47**A). Localized actin polymerization, initiated by Rho-family GTPases and their activating Rho-GEFs (discussed in Chapters 15 and 16), shapes the pseudopods. An active Rho GTPase switches on the kinase activity of local PI kinases, and initial actin polymerization occurs in response to an accumulation of $PI(4,5)P_2$ in the membrane (see Figure 13–11). To seal off the phagosome and complete its engulfment, actin is depolymerized at its base as $PI(4,5)P_2$ is subjected to a PI 3-kinase, which converts it to $PI(3,4,5)P_3$. $PI(3,4,5)P_3$ is required for closure of the phagosome and may also contribute to reshaping the actin network to help drive the invagination of the forming phagosome (Figure 13–47B). In this way, the ordered generation and consumption of specific phosphoinositides guides sequential steps in phagosome formation.

Several other classes of receptors that promote phagocytosis have been characterized. Some recognize *complement* components, which collaborate with antibodies in targeting microbes for destruction (discussed in Chapter 24). Others directly recognize oligosaccharides on the surface of certain microorganisms. Still others recognize cells that have died by apoptosis. Apoptotic cells lose the asymmetric distribution of phospholipids in their plasma membrane. As a consequence, negatively charged phosphatidylserine, which is normally confined to the cytosolic leaflet of the lipid bilayer, is now exposed on the outside of the cell, where it helps to trigger the phagocytosis of the dead cell.

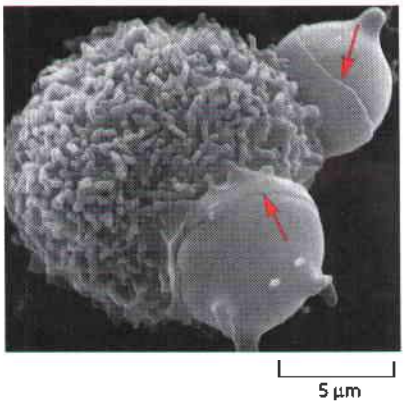


Figure 13–46 Phagocytosis by a macrophage. A scanning electron micrograph of a mouse macrophage phagocytosing two chemically altered red blood cells. The *red arrows* point to edges of thin processes (pseudopods) of the macrophage that are extending as collars to engulf the red cells. (Courtesy of Jean Paul Revel.)

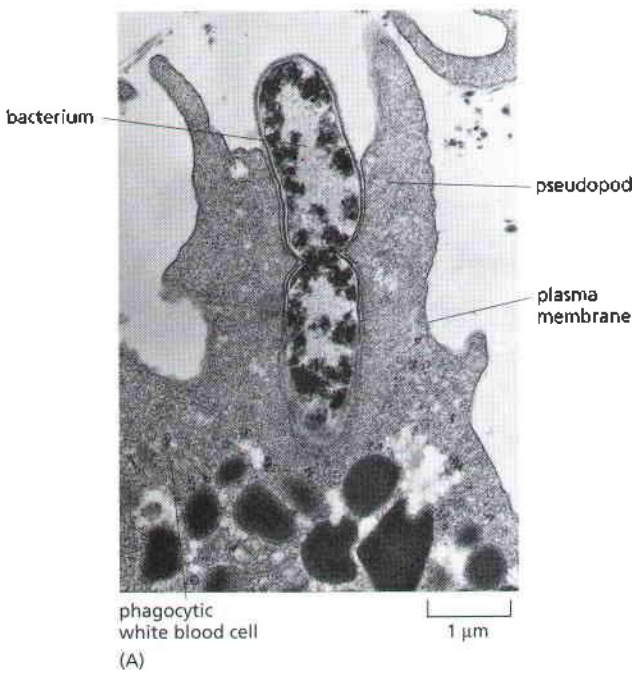


Figure 13–47 A neutrophil reshaping a plasma membrane during phagocytosis. (A) An electron micrograph of a neutrophil phagocytosing a bacterium, which is in the process of dividing. (B) Pseudopod extension and phagosome formation are driven by actin polymerization and reorganization, which respond to the accumulation of specific phosphoinositides in the membrane of the forming phagosome. (A, courtesy of Dorothy F. Bainton, Phagocytic Mechanisms in Health and Disease. New York: Intercontinental Medical Book Corporation, 1971.)

Remarkably, macrophages will also phagocytose a variety of inanimate particles—such as glass or latex beads and asbestos fibers—yet they do not phagocytose live animal cells. Living animal cells seem to display "don't-eat-me" signals in the form of cell-surface proteins that bind to inhibiting receptors on the surface of macrophages. The inhibitory receptors recruit tyrosine phosphatases that antagonize the intracellular signaling events required to initiate phagocytosis, thereby locally inhibiting the phagocytic process. Thus phagocytosis, like many other cell processes, depends on a balance between positive signals that activate the process and negative signals that inhibit it. Apopototic cells are thought both to gain "eat-me" signals (such as extracellularly exposed phosphatidylserine) and to lose their "don't-eat-me" signals, causing them to be very rapidly phagocytosed by macrophages.

Pinocytic Vesicles Form from Coated Pits in the Plasma Membrane

Virtually all eucaryotic cells continually ingest bits of their plasma membrane in the form of small pinocytic (endocytic) vesicles, which are later returned to the cell surface. The rate at which plasma membrane is internalized in this process of **pinocytosis** varies between cell types, but it is usually surprisingly large. A macrophage, for example, ingests 25% of its own volume of fluid each hour. This means that it must ingest 3% of its plasma membrane each minute, or 100% in about half an hour. Fibroblasts endocytose at a somewhat lower rate (1% of their plasma membrane per minute), whereas some amoebae ingest their plasma membrane even more rapidly. Since a cell's surface area and volume remain unchanged during this process, it is clear that the same amount of membrane being removed by endocytosis is being added to the cell surface by the converse process of *exocytosis*. In this sense, endocytosis and exocytosis are linked processes that can be considered to constitute an *endocytic–exocytic cycle*. The coupling between exocytosis and endocytosis is particularly strict in specialized structures characterized by high membrane turnover, such as the neuronal synapse.

The endocytic part of the cycle often begins at **clathrin-coated pits**. These specialized regions typically occupy about 2% of the total plasma membrane area. The lifetime of a clathrin-coated pit is short: within a minute or so of being formed, it invaginates into the cell and pinches off to form a clathrin-coated vesicle (**Figure 13–48**). It has been estimated that about 2500 clathrin-coated vesicles leave the plasma membrane of a cultured fibroblast every minute. The coated vesicles are even more transient than the coated pits: within seconds of being formed, they shed their coat and are able to fuse with early endosomes. Since extracellular fluid is trapped in clathrin-coated pits as they invaginate to

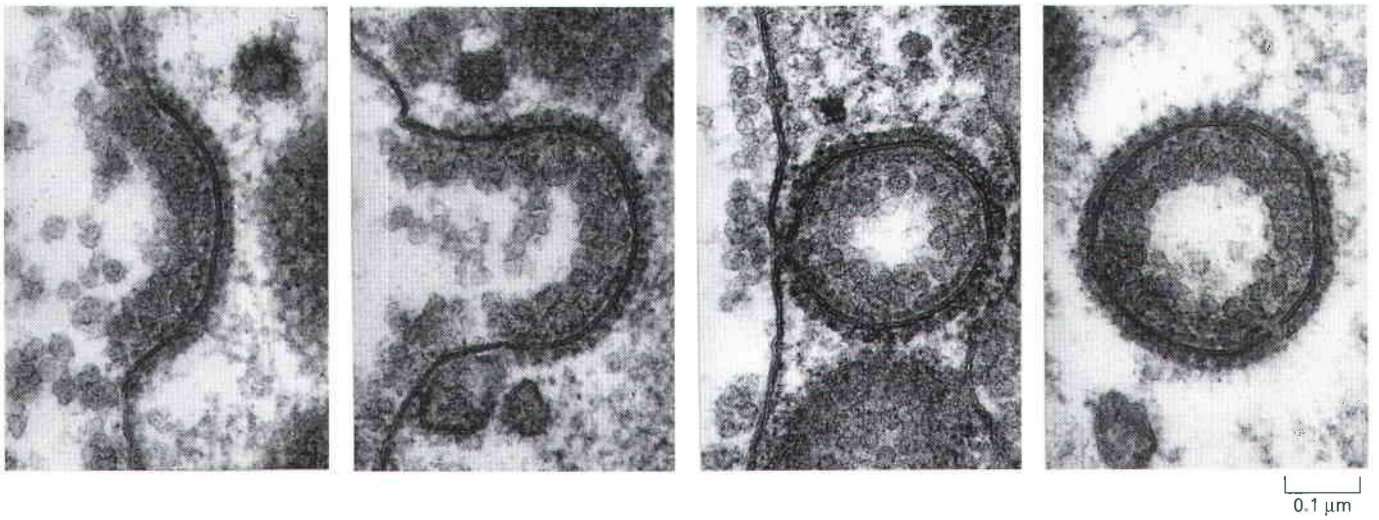


Figure 13–48 The formation of clathrin-coated vesicles from the plasma membrane. These electron micrographs illustrate the probable sequence of events in the formation of a clathrin-coated vesicle from a clathrin-coated pit. The clathrin-coated pits and vesicles shown are larger than those seen in normal-sized cells. They take up lipoprotein particles into a very large hen oocyte to form yolk. The lipoprotein particles bound to their membrane-bound receptors appear as a dense, fuzzy layer on the extracellular surface of the plasma membrane—which is the inside surface of the vesicle. (Courtesy of M.M. Perry and A.B. Gilbert, *J. Cell Sci.* 39:257–272, 1979. With permission from The Company of Biologists.)

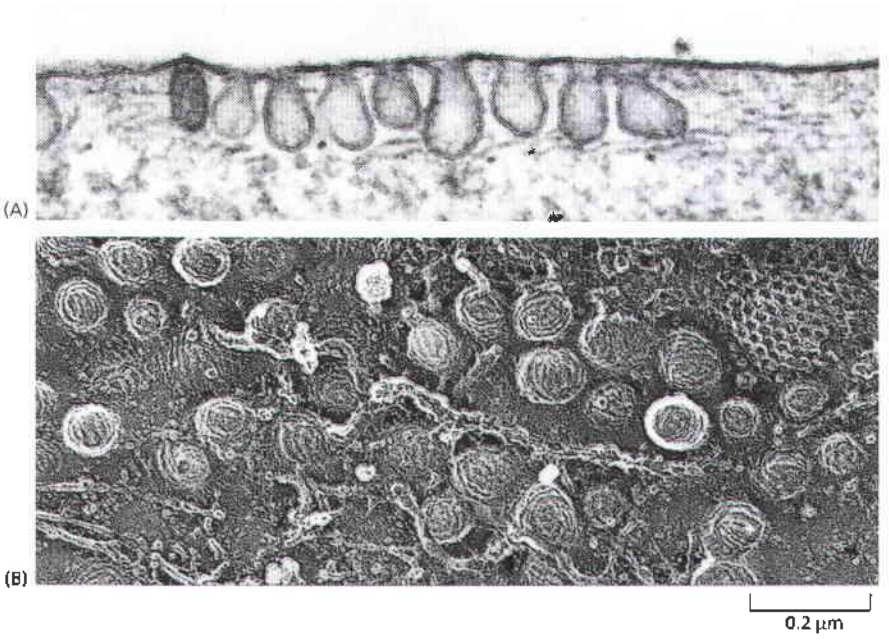


Figure 13–49 Caveolae in the plasma membrane of a fibroblast. (A) This electron micrograph shows a plasma membrane with a very high density of caveolae. Note that no cytosolic coat is visible. (B) This rapid-freeze deep-etch image demonstrates the characteristic "cauliflower" texture of the cytosolic face of the caveolae membrane. The regular texture is thought to result from aggregates of caveolin in the membrane. A clathrin-coated pit is also seen at the upper right. (Courtesy of R.G.W. Anderson, from K.G. Rothberg et al., *Cell* 68:673–682, 1992. With permission from Elsevier.)

form coated vesicles, any substance dissolved in the extracellular fluid is internalized—a process called *fluid-phase endocytosis.*

Not All Pinocytic Vesicles Are Clathrin-Coated

In addition to clathrin-coated pits and vesicles, there are other, less well-understood mechanisms by which cells can form pinocytic vesicles. One of these pathways initiates at **caveolae** (from the Latin for "little cavities"), originally recognized by their ability to transport molecules across endothelial cells, which form the inner lining of blood vessels. Caveolae are present in the plasma membrane of most cell types, and in some of these they are seen in the electron microscope as deeply invaginated flasks (**Figure 13–49**). They are thought to form from membrane microdomains, or *lipid rafts*, which are patches of the plasma membrane that are especially rich in cholesterol, glycosphingolipids, and GPI-anchored membrane proteins (see Figure 10–14). The major structural proteins in caveolae are **caveolins**, which are a family of unusual integral membrane proteins that each insert a hydrophobic loop into the membrane from the cytosolic side but do not extend across the membrane.

In contrast to clathrin-coated and COPI- or COPII-coated vesicles, caveolae are thought to invaginate and collect cargo proteins by virtue of the lipid composition of the calveolar membrane, rather than by the assembly of a cytosolic protein coat. Caveolins may stabilize these raft domains, into which certain plasma membrane proteins partition. Caveolae pinch off from the plasma membrane using dynamin, and they deliver their contents either to an endosome-like compartment (called a *caveosome*) or to the plasma membrane on the opposite side of a polarized cell (in a process called *transcytosis*, which we discuss later). Because caveolins are integral membrane proteins, they do not dissociate from the vesicles after endocytosis; instead they are delivered to the target compartments, where they are maintained as discrete membrane domains. Some animal viruses such as SV40 and papilloma virus (which causes warts) enter cells in vesicles derived from caveolae. The viruses are first delivered to caveosomes, and they move from there in specialized transport vesicles to the ER. The viral genome exits from the ER across the ER membrane into the cytosol, from where it is imported into the nucleus to start the infection cycle.

Endocytic vesicles can also bud from caveolin-free raft domains on the plasma membrane and deliver their cargo to caveosomes. Molecules that enter the cell through caveosomes avoid endosomes and lysosomes and are therefore shielded from exposure to low pH and lysosomal hydrolases; it is unknown how they move from caveosomes to other destinations in the cell.

Cells Use Receptor-Mediated Endocytosis to Import Selected Extracellular Macromolecules

In most animal cells, clathrin-coated pits and vesicles provide an efficient pathway for taking up specific macromolecules from the extracellular fluid. In this process, called **receptor-mediated endocytosis**, the macromolecules bind to complementary transmembrane receptor proteins, accumulate in coated pits, and then enter the cell as receptor–macromolecule complexes in clathrin-coated vesicles (see Figure 13–48). Because ligands are selectively captured by receptors, receptor-mediated endocytosis provides a selective concentrating mechanism that increases the efficiency of internalization of particular ligands more than a hundredfold. In this way, even minor components of the extracellular fluid can be specifically taken up in large amounts without taking in a large volume of extracellular fluid. A particularly well-understood and physiologically important example is the process that mammalian cells use to take up cholesterol.

Many animals cells take up cholesterol through receptor-mediated endocytosis and, in this way, acquire most of the cholesterol they require to make new membrane. If the uptake is blocked, cholesterol accumulates in the blood and can contribute to the formation in blood vessel (artery) walls of *atherosclerotic plaques*, deposits of lipid and fibrous tissue that can cause strokes and heart attacks by blocking arterial blood flow. In fact, it was a study of humans with a strong genetic predisposition for *atherosclerosis* that first revealed the mechanism of receptor-mediated endocytosis.

Most cholesterol is transported in the blood as cholesteryl esters in the form of lipid–protein particles known as **low-density lipoproteins** (**LDLs**) (**Figure 13–50**). When a cell needs cholesterol for membrane synthesis, it makes transmembrane receptor proteins for LDL and inserts them into its plasma membrane. Once in the plasma membrane, the *LDL receptors* diffuse until they associate with clathrin-coated pits that are in the process of forming (**Figure 13–51**A). Since coated pits constantly pinch off to form coated vesicles, any LDL particles bound to LDL receptors in the coated pits are rapidly internalized in coated vesicles. After shedding their clathrin coats, the vesicles deliver their contents to early endosomes, which are located near the cell periphery. Once the LDL and LDL receptors encounter the low pH in the endosomes, LDL is released from its receptor and is delivered via late endosomes to lysosomes. There, the cholesteryl esters in the LDL particles are hydrolyzed to free cholesterol, which is now available to the cell for new membrane synthesis. If too much free cholesterol accumulates in a cell, the cell shuts off both its own cholesterol synthesis and the synthesis of LDL receptors, so that it ceases either to make or to take up cholesterol.

This regulated pathway for cholesterol uptake is disrupted in individuals who inherit defective genes encoding LDL receptors. The resulting high levels of blood cholesterol predispose these individuals to develop atherosclerosis

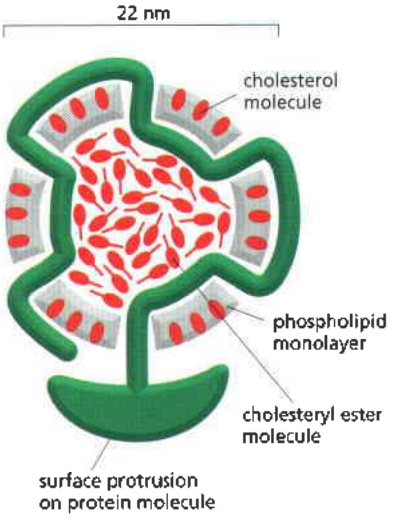


Figure 13–50 A low-density lipoprotein (LDL) particle. Each spherical particle has a mass of 3×10^6 daltons. It contains a core of about 1500 cholesterol molecules esterified to long-chain fatty acids. A lipid monolayer composed of about 800 phospholipid and 500 unesterified cholesterol molecules surrounds the core of cholesterol esters. A single molecule of a 500,000-dalton protein organizes the particle and mediates the specific binding of LDL to cell-surface LDL receptors.

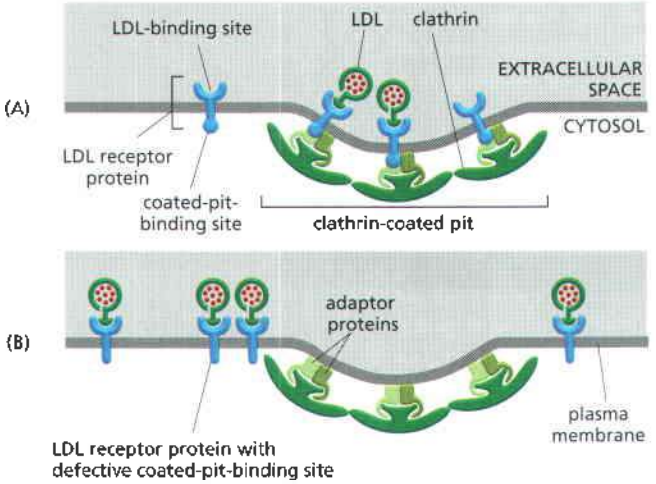


Figure 13–51 Normal and mutant LDL receptors. (A) LDL receptors binding to a coated pit in the plasma membrane of a normal cell. The human LDL receptor is a single-pass transmembrane glycoprotein composed of about 840 amino acids, only 50 of which are on the cytoplasmic side of the membrane. (B) A mutant cell in which the LDL receptors are abnormal and lack the site in the cytoplasmic domain that enables them to bind to adaptor proteins in the clathrin-coated pits. Such cells bind LDL but cannot ingest it. In most human populations, 1 in 500 individuals inherits one defective LDL receptor gene and, as a result, has an increased risk of a heart attack caused by atherosclerosis.

prematurely, and many would die at an early age of heart attacks resulting from coronary artery disease if they were not treated with drugs that lower the level of blood cholesterol. In some cases, the receptor is lacking altogether. In others, the receptors are defective—in either the extracellular binding site for LDL or the intracellular binding site that attaches the receptor to the coat of a clathrin-coated pit (see Figure 13–51B). In the latter case, normal numbers of LDL receptors are present, but they fail to become localized in clathrin-coated pits. Although LDL binds to the surface of these mutant cells, it is not internalized, directly demonstrating the importance of clathrin-coated pits for the receptor-mediated endocytosis of cholesterol.

More than 25 distinct receptors are known to participate in receptor-mediated endocytosis of different types of molecules. They all apparently use clathrin-dependent internalization routes and are guided into clathrin-coated pits by signals in their cytoplasmic tails that bind to adaptor proteins in the clathrin coat. Many of these receptors, like the LDL receptor, enter coated pits irrespective of whether they have bound their specific ligands. Others enter preferentially when bound to a specific ligand, suggesting that a ligand-induced conformational change is required for them to activate the signal sequence that guides them into the pits. Since most plasma membrane proteins fail to become concentrated in clathrin-coated pits, the pits serve as molecular filters, preferentially collecting certain plasma membrane proteins (receptors) over others.

Electron-microscope studies of cultured cells exposed simultaneously to different labeled ligands demonstrate that many kinds of receptors can cluster in the same coated pit, whereas some other receptors cluster in different clathrin-coated pits. The plasma membrane of one clathrin-coated pit can probably accommodate up to 1000 receptors of assorted varieties. Although all of the receptor–ligand complexes that use this endocytic pathway are apparently delivered to the same endosomal compartment, the subsequent fates of the endocytosed molecules vary, as we discuss next.

Endocytosed Materials That Are Not Retrieved from Endosomes End Up in Lysosomes

The endosomal compartments of a cell can be complex. They can be made visible in the electron microscope by adding a readily detectable tracer molecule, such as the enzyme peroxidase, to the extracellular medium and leaving the cells for various lengths of time to take it up by endocytosis. The distribution of the molecule after its uptake reveals the endosomal compartments as a set of heterogeneous, membrane-enclosed tubes extending from the periphery of the cell to the perinuclear region, where it is often close to the Golgi apparatus. Two sequential sets of endosomes can be readily distinguished in such labeling experiments. The tracer molecule appears within a minute or so in **early endosomes**, just beneath the plasma membrane. After 5–15 minutes, it has moved to **late endosomes**, close to the Golgi apparatus and near the nucleus. Early and late endosomes differ in their protein compositions. The transition from early to late endosomes is accompanied by the release of Rab5 and the binding of Rab7, for example.

As mentioned earlier, a vacuolar H^+ ATPase in the endosomal membrane, which pumps H^+ into endosomes from the cytosol, keeps the lumen of the endosomal compartments acidic (pH ~6). In general, later endosomes are more acidic than early endosomes. This gradient of acidic environments has a crucial role in the function of these organelles.

We have already seen how endocytosed materials mix in early endosomes with newly synthesized acid hydrolases and eventually end up being degraded in lysosomes. Many molecules, however, are specifically diverted from this journey to destruction. They are instead recycled from the early endosomes back to the plasma membrane via transport vesicles. Only molecules that are not retrieved from endosomes in this way are delivered to lysosomes for degradation.

Although mild digestion may start in early endosomes, many hydrolases are synthesized and delivered there as proenzymes, called *zymogens*, which contain extra inhibitory domains at their N-terminus that keep the hydrolase inactive until these domains are proteolytically removed. The hydrolases are activated when late endosomes become endolysosomes as the result of fusion with pre-existing lysosomes, which contain a full complement of active hydrolases that digest off the inhibitory domains from the newly synthesized enzymes. Moreover, the pH in early endosomes is not low enough to activate lysosomal hydrolases optimally. By these means, cells can retrieve most membrane proteins from early endosomes and recycle them back to the plasma membrane.

Specific Proteins Are Retrieved from Early Endosomes and Returned to the Plasma Membrane

Early endosomes form a compartment that acts as the main sorting station in the endocytic pathway, just as the *cis* and *trans* Golgi networks serve this function in the biosynthetic–secretory pathway. In the mildly acidic environment of the early endosome, many internalized receptor proteins change their conformation and release their ligand, as already discussed for the M6P receptors. Those endocytosed ligands that dissociate from their receptors in the early endosome are usually doomed to destruction in lysosomes, along with the other soluble contents of the endosome. Some other endocytosed ligands, however, remain bound to their receptors, and thereby share the fate of the receptors.

The fates of receptors—and of any ligands remaining bound to them—vary according to the specific type of receptor. (1) Most receptors are recycled and return to the same plasma membrane domain from which they came; (2) some proceed to a different domain of the plasma membrane, thereby mediating *transcytosis*; and (3) some progress to lysosomes, where they are degraded (**Figure 13–52**).

The LDL receptor follows the first pathway. It dissociates from its ligand, LDL, in the early endosome and is recycled back to the plasma membrane for reuse, leaving the discharged LDL to be carried to lysosomes (**Figure 13–53**). The recycling transport vesicles bud from long, narrow tubules that extend from the

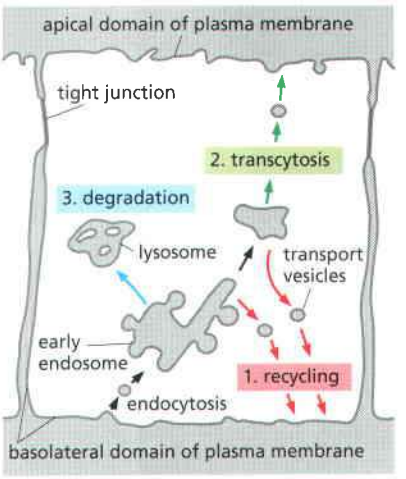


Figure 13–52 Possible fates for transmembrane receptor proteins that have been endocytosed. Three pathways from the endosomal compartment in an epithelial cell are shown. Retrieved receptors are returned (1) to the same plasma membrane domain from which they came *(recycling)* or (2) to a different domain of the plasma membrane *(transcytosis)*. (3) Receptors that are not specifically retrieved from endosomes follow the pathway from the endosomal compartment to lysosomes, where they are degraded *(degradation)*. The formation of oligomeric aggregates in the endosomal membrane may be one of the signals that guide receptors into the degradative pathway. If the ligand that is endocytosed with its receptor stays bound to the receptor in the acidic environment of the endosome, it follows the same pathway as the receptor; otherwise it is delivered to lysosomes.

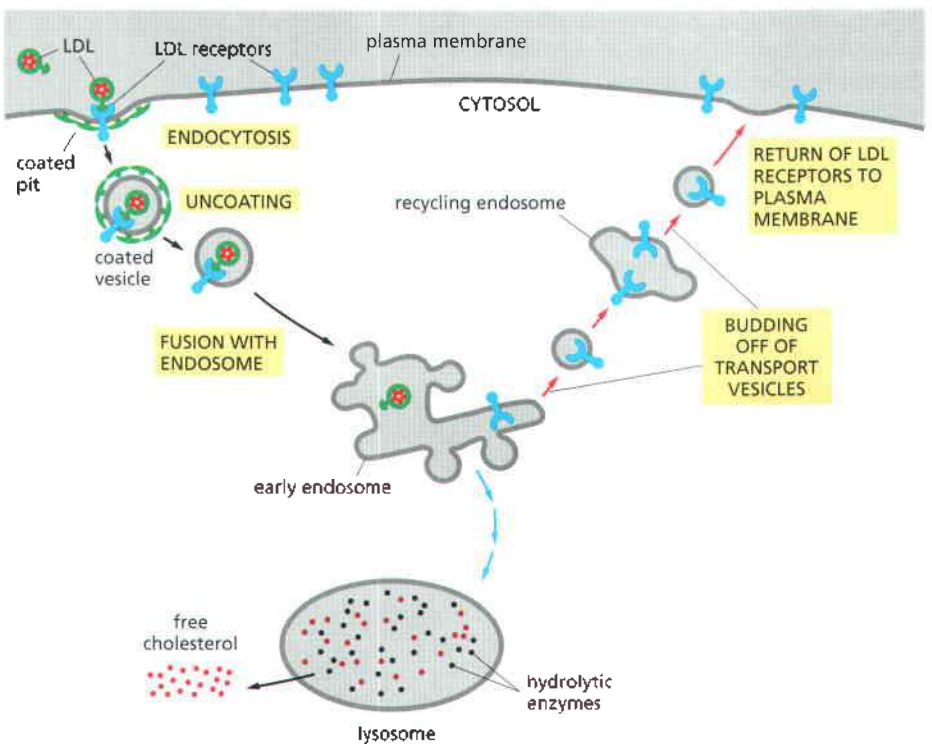


Figure 13–53 The receptor-mediated endocytosis of LDL. <GCTA> Note that the LDL dissociates from its receptors in the acidic environment of the early endosome. After a number of steps (shown in Figure 13–55), the LDL ends up in lysosomes, where it is degraded to release free cholesterol. In contrast, the LDL receptors are returned to the plasma membrane via clathrin-coated transport vesicles that bud off from the tubular region of the early endosome, as shown. For simplicity, only one LDL receptor is shown entering the cell and returning to the plasma membrane. Whether it is occupied or not, an LDL receptor typically makes one round trip into the cell and back to the plasma membrane every 10 minutes, making a total of several hundred trips in its 20-hour lifespan.

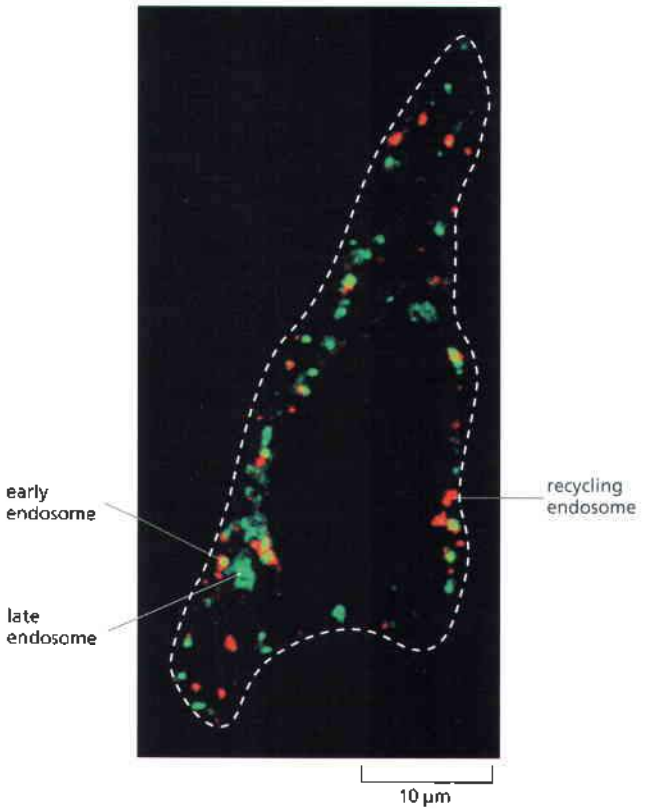


Figure 13–54 Sorting of membrane proteins in the endocytic pathway. Transferrin receptors mediate iron uptake and constitutively cycle between endosomes and the plasma membrane. By contrast, activated opioid receptors are down-regulated by endocytosis followed by degradation in lysosomes; they are activated by opiates such as morphine and heroin, as well as by endogenous peptides called enkephalins and endorphins. Endocytosis of both types of receptors starts in clathrin-coated pits. The receptors are then delivered to early endosomes, where their pathways part: transferrin receptors are sorted to the recycling endosomes, whereas opioid receptors are sorted to late endosomes. The micrograph shows both receptors—labeled with different fluorescent dyes—30 min after endocytosis (transferrin receptors are labeled in *red* and opioid receptors in *green*). At this time, some early endosomes still contain both receptors and are seen as *yellow*, due to the overlap of red and green light emitted from the fluorescent dyes. By contrast, recycling endosomes and late endosomes are selectively enriched in either transferrin or opioid receptors, respectively, and therefore appear as distinct red and green structures. (Courtesy of Mark von Zastrow.)

early endosomes. It is likely that the geometry of these tubules helps the sorting process: because tubules have a large membrane area enclosing a small volume, membrane proteins become enriched over soluble proteins. Transport vesicles that return material to the plasma membrane begin budding from the tubules, but tubular portions of the early endosome also pinch off and fuse with one another to form *recycling endosomes*, which serve as way-stations for the traffic between early endosomes and the plasma membrane. This recycling pathway operates continuously, compensating for the continuous endocytosis occurring at the plasma membrane.

The **transferrin receptor** follows a similar recycling pathway as the LDL receptor, but unlike the LDL receptor it also recycles its ligand. Transferrin is a soluble protein that carries iron in the blood. Cell-surface transferrin receptors deliver transferrin with its bound iron to early endosomes by receptor-mediated endocytosis. The low pH in the endosome induces transferrin to release its bound iron, but the iron-free transferrin itself (called apotransferrin) remains bound to its receptor. The receptor–apotransferrin complex enters the tubular extensions of the early endosome and from there is recycled back to the plasma membrane (**Figure 13–54**). When the apotransferrin returns to the neutral pH of the extracellular fluid, it dissociates from the receptor and is thereby freed to pick up more iron and begin the cycle again. Thus, transferrin shuttles back and forth between the extracellular fluid and the endosomal compartment, avoiding lysosomes and delivering iron to the cell interior, as needed for cells to grow and proliferate.

The second pathway that endocytosed receptors can follow from endosomes is taken by many signaling receptors, including opioid receptors (see Figure 13–54) and the receptor that binds *epidermal growth factor (EGF)*. EGF is a small, extracellular signal protein that stimulates epidermal and various other cells to divide. Unlike LDL receptors, EGF receptors accumulate in clathrin-coated pits only after binding EGF, and most of them do not recycle but are degraded in lysosomes, along with the ingested EGF. EGF binding therefore first activates intracellular signaling pathways and then leads to a decrease in the concentration of EGF receptors on the cell surface, a process called *receptor down-regulation* that reduces the cell's subsequent sensitivity to EGF (see Figure 15–29).

Clathrin-dependent receptor-mediated endocytosis is highly regulated. The receptors are first covalently modified with the small protein ubiquitin. But, unlike *polyubiquitylation*, which adds a chain of ubiquitins that typically targets a protein for degradation in proteasomes (discussed in Chapter 6), ubiquitin tagging for sorting into the clathrin-dependent endocytic pathway adds one or more single ubiquitin molecules to the protein—a process called *monoubiquitylation* or *multiubiquitylation*, respectively. Ubiquitin-binding proteins recognize the attached ubiquitin and help direct the modified receptors into clathrin-coated pits. After delivery to the endosome, other ubiquitin-binding proteins recognize the ubiquitin and help mediate sorting steps.

Multivesicular Bodies Form on the Pathway to Late Endosomes

As previously stated, many of the endocytosed molecules move from the early to the late endosomal compartment. In this process, early endosomes migrate slowly along microtubules toward the cell interior, while shedding membrane tubules and vesicles that recycle material to the plasma membrane and TGN. At the same time, the membrane enclosing the migrating endosomes forms invaginating buds that pinch off and form internal vesicles; they are then called **multivesicular bodies** (Figure 13–55). Multivesicular bodies eventually fuse with a late endosomal compartment or with each other to become late endosomes. At the end of this pathway, the late endosomes convert to endolysosomes and lysosomes as a result of both their fusion with preexisting lysosomes and progressive acidification (Figure 13–56).

The multivesicular bodies carry those endocytosed membrane proteins that are to be degraded. As part of the protein-sorting process, receptors destined for degradation, such as occupied EGF receptors described previously, selectively partition into the invaginating membrane of the multivesicular bodies. In this way, both the receptors and any signaling proteins strongly bound to them are made fully accessible to the digestive enzymes that will degrade them (Figure 13–57). In addition to endocytosed membrane proteins, multivesicular bodies also contain the soluble content of early endosomes destined for late endosomes and digestion in lysosomes.

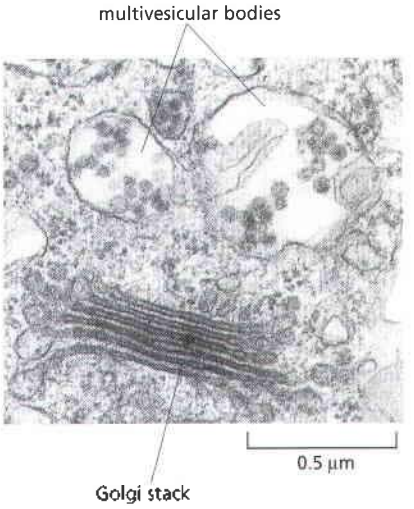


Figure 13–55 **Electron micrograph of a multivesicular body in a plant cell.** The large amount of internal membrane will be delivered to the vacuole, the plant equivalent of the lysosome, for digestion.

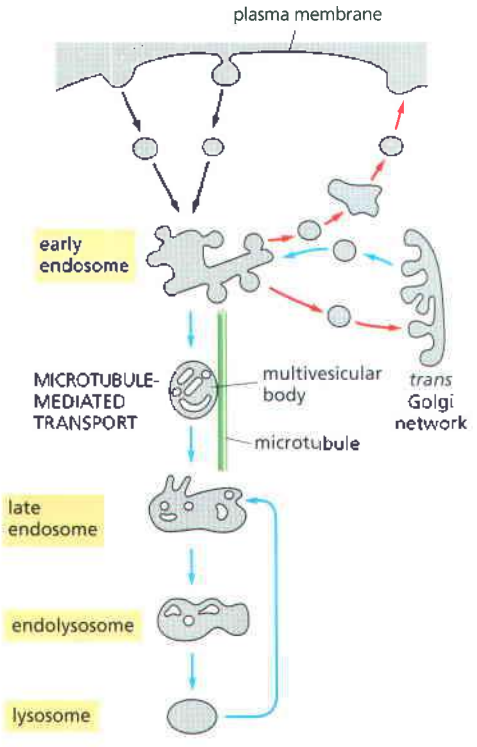


Figure 13–56 **Details of the endocytic pathway from the plasma membrane to lysosomes.** Maturation of early endosomes to late endosomes occurs through the formation of multivesicular bodies, which contain large amounts of invaginated membrane and internal vesicles (hence their name). Multivesicular bodies move inward along microtubules, continually shedding transport vesicles that recycle components to the plasma membrane. They gradually convert into late endosomes, either by fusing with each other or by fusing with preexisting late endosomes. The late endosomes no longer send vesicles to the plasma membrane.

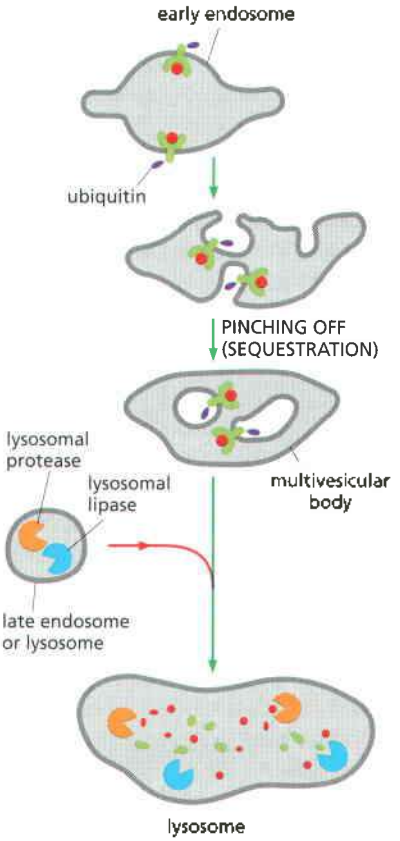


Figure 13–57 The sequestration of endocytosed proteins into internal membranes of multivesicular bodies. Eventually, proteases and lipases in lysosomes digest all of the internal membranes within multivesicular bodies produced by the invaginations. The invagination processes are essential to achieve complete digestion of endocytosed membrane proteins: because the outer membrane of the multivesicular body becomes continuous with the lysosomal membrane, for example, lysosomal hydrolases could not digest the cytosolic domains of endocytosed transmembrane proteins such as the EGF receptor shown here, if the protein were not localized in internal vesicles.

Sorting into the internal vesicles of a multivesicular body requires one or multiple ubiquitin tags, which are added to the cytosolic domains of membrane proteins. These tags initially help guide the proteins into clathrin-coated vesicles. Once delivered to the endosomal membrane, the ubiquitin tags are recognized again, this time by a series of cytosolic protein complexes, called *ESCRT-0, -I, -II,* and *-III,* which bind sequentially, handing the ubiquitylated cargo from one complex to the next, and ultimately mediate the sorting process into the internal vesicles of multivesicular bodies (**Figure 13–58**). Membrane invagination into multivesicular bodies also depends on a lipid kinase that phosphorylates phosphatidylinositol to produce PI(3)P, which serves as an additional docking site for the ESCRT complexes; these complexes require both PI(3)P and the presence of ubiquitylated cargo proteins to bind to the endosomal membrane. A second PI kinase adds another phosphate group to PI(3)P, producing $PI(3,5)P_2$, which is required for ESCRT-III to form large multimeric assemblies on the membrane. It is not known how the assembly of ESCRT complexes ultimately

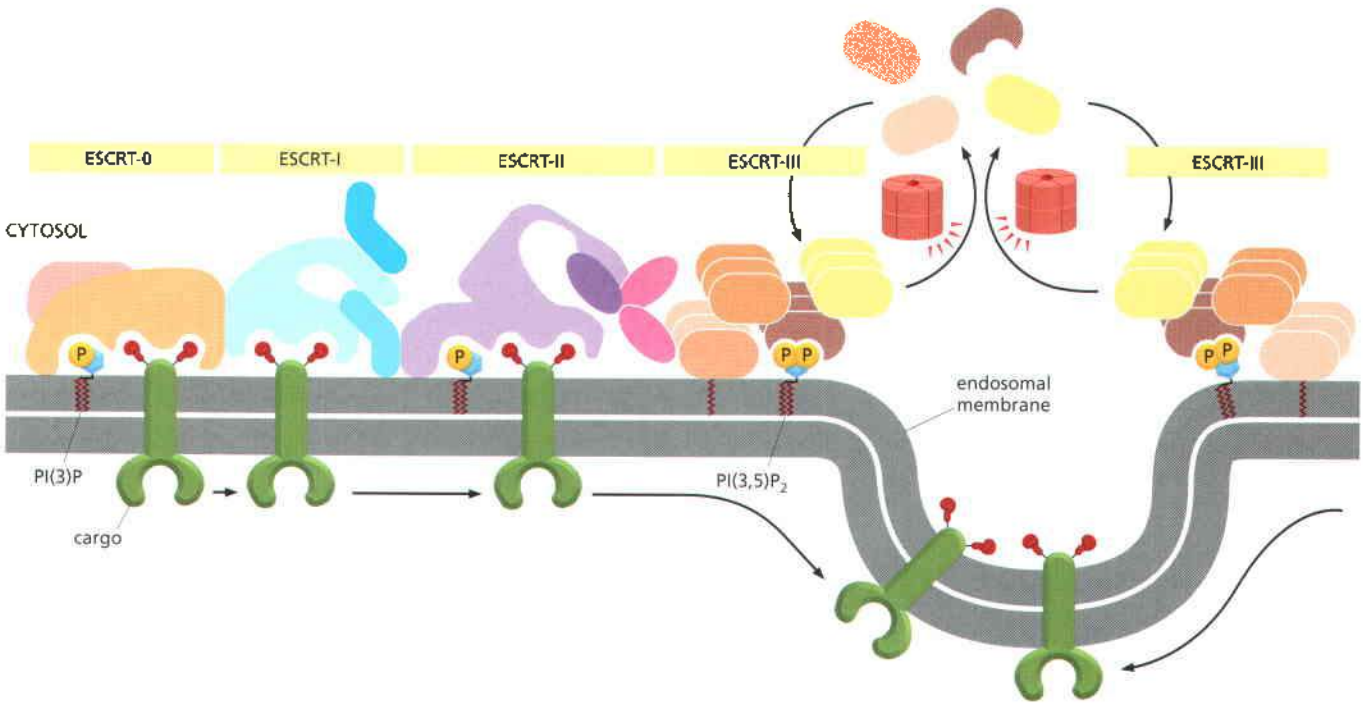


Figure 13–58 Sorting of endocytosed membrane proteins into the internal vesicles of a multivesicular body. A series of complex binding events passes the ubiquitylated cargo proteins sequentially from one ESCRT complex to the next, eventually concentrating them in membrane areas that bud away from the cytosol into the lumen of the endosome to form the internal membrane vesicles of the multivesicular body. ESCRT complexes are soluble in the cytosol and are recruited to the membrane as needed. First, ESCRT-0 binds both the ubiquitin attached to the cargo protein and to PI(3)P head groups. ESCRT-0 dissociates from the membrane, handing the ubiquitylated cargo protein over to the ESCRT-I complex; next ESCRT-I dissociates, handing the cargo protein over to ESCRT-II complex; and finally ESCRT-II dissociates and ESCRT-III complexes assemble on the membrane. By contrast to ESCRT-0, -I, and –II, ESCRT-III does not bind to the ubiquitylated cargo directly. Instead its assembly into expansive multimeric structures is thought to confine the cargo molecules into specialized membrane areas that then invaginate, leaving the ESCRT components on the endosome surface. An AAA-ATPase (*red* cylinders) then disassembles the ESCRT-III complexes so that they can be reused.

drives the invagination and pinching-off processes required to form the internal vesicles but the ESCRT complexes themselves are not part of the invaginating membranes.

Mutant cells compromised in ESCRT function display signaling defects. In such cells, activated receptors cannot be down-regulated by endocytosis and packaged into multivesicular bodies and therefore mediate prolonged signaling, which can lead to uncontrolled cell proliferation and cancer.

The same ESCRT machinery that drives the internal budding from the endosomal membrane to form multivesicular bodies is also used by HIV, ebola, and other enveloped viruses to bud from the plasma membrane into the extracellular space. The two processes are topologically equivalent, as they both involve budding away from the cytosolic surface of the membrane (Figure 13–59).

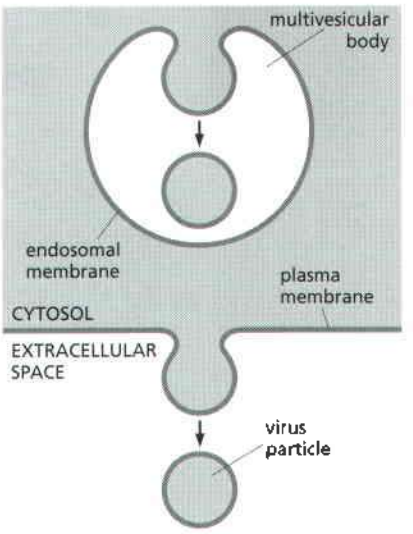


Figure 13–59 **ESCRT complexes in multivesicular body formation and virus budding.** In the two topologically equivalent processes indicated by the arrows, ESCRT complexes shape membranes into buds that bulge away from the cytosol.

Transcytosis Transfers Macromolecules Across Epithelial Cell Sheets

Some receptors on the surface of polarized epithelial cells transfer specific macromolecules from one extracellular space to another by **transcytosis** (Figure 13–60). These receptors are endocytosed and then follow a pathway from endosomes to a different plasma membrane domain (see Figure 13–52). A newborn rat, for example, obtains antibodies from its mother's milk (which help protect it against infection) by transporting them across the epithelium of its gut. The lumen of the gut is acidic, and, at this low pH, the antibodies in the milk bind to specific receptors on the apical (absorptive) surface of the gut epithelial cells. The receptor–antibody complexes are internalized via clathrin-coated pits and vesicles and are delivered to early endosomes. The complexes remain intact and are retrieved in transport vesicles that bud from the early

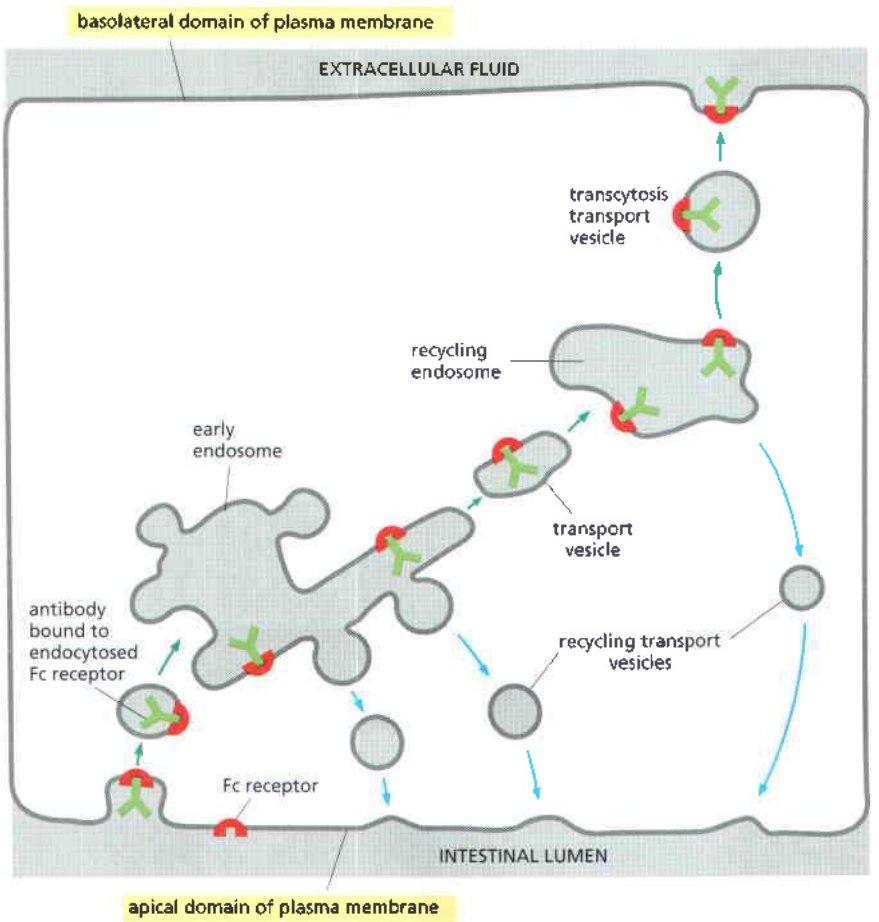


Figure 13–60 **The role of recycling endosomes in transcytosis.** Recycling endosomes form a way-station on the transcytotic pathway. In the example shown here, an antibody receptor on a gut epithelial cell binds antibody and is endocytosed, eventually carrying the antibody to the basolateral plasma membrane. The receptor is called an Fc receptor because it binds the Fc part of the antibody (discussed in Chapter 25).

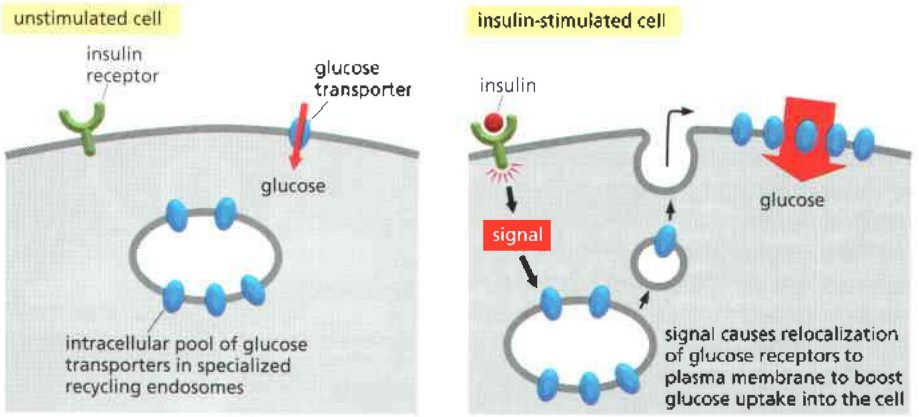


Figure 13–61 **Storage of plasma membrane proteins in recycling endosomes.** Recycling endosomes can serve as an intracellular pool of specialized plasma membrane proteins that can be mobilized when needed. In the example shown, insulin binding to the insulin receptor triggers an intracellular signaling pathway that causes the rapid insertion of glucose transporters into the plasma membrane of a fat or muscle cell, greatly increasing glucose intake.

endosome and subsequently fuse with the basolateral domain of the plasma membrane. On exposure to the neutral pH of the extracellular fluid that bathes the basolateral surface of the cells, the antibodies dissociate from their receptors and eventually enter the newborn's bloodstream.

The transcytotic pathway from the early endosome to the plasma membrane is not direct. The receptors first move from the early endosome to an intermediate endosomal compartment, the **recycling endosome** described previously (see Figure 13–60). The variety of pathways that different receptors follow from endosomes implies that, in addition to binding sites for their ligands and binding sites for coated pits, many receptors also possess sorting signals that guide them into the appropriate type of transport vesicle leaving the endosome and moving to the appropriate target membrane in the cell.

A unique property of recycling endosomes is that cells can regulate the exit of membrane proteins from the compartment. Thus, cells can adjust the flux of proteins through the transcytotic pathway according to need. Although the mechanism is uncertain, this regulation allows recycling endosomes to play an important part in adjusting the concentration of specific plasma membrane proteins. Fat cells and muscle cells, for example, contain large intracellular pools of the glucose transporters that are responsible for the uptake of glucose across the plasma membrane. These membrane transport proteins are stored in specialized recycling endosomes until the hormone *insulin* stimulates the cell to increase its rate of glucose uptake. In response to the insulin signal, transport vesicles rapidly bud from the recycling endosome and deliver large numbers of glucose transporters to the plasma membrane, thereby greatly increasing the rate of glucose uptake into the cell (**Figure 13–61**).

Epithelial Cells Have Two Distinct Early Endosomal Compartments but a Common Late Endosomal Compartment

In polarized epithelial cells, endocytosis occurs from both the *basolateral domain* and the *apical domain* of the plasma membrane. Material endocytosed from either domain first enters an early endosomal compartment that is unique to that domain. This arrangement allows endocytosed receptors to be recycled back to their original membrane domain, unless they contain sorting signals that mark them for transcytosis to the other domain. Molecules endocytosed from either plasma membrane domain that are not retrieved from the early endosomes end up in a common late endosomal compartment near the cell center and are eventually degraded in lysosomes (**Figure 13–62**).

Whether cells contain a few connected or many unconnected endosomal compartments seems to depend on the cell type and the physiological state of the cell. Like many other membrane-enclosed organelles, endosomes of the same type can readily fuse with one another (an example of homotypic fusion, discussed earlier) to create large continuous endosomes.

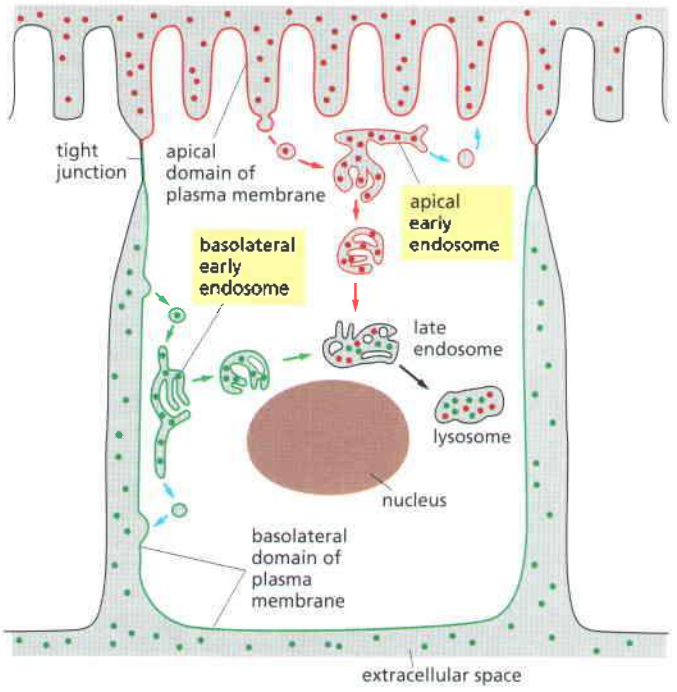


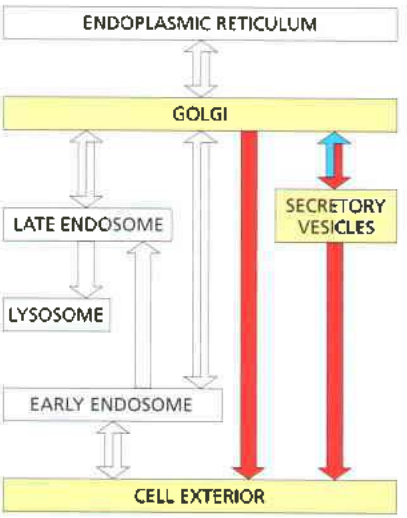
Figure 13–62 The two distinct early endosomal compartments in an epithelial cell. The basolateral and the apical domains of the plasma membrane communicate with separate early endosomal compartments. However, endocytosed molecules from both domains that do not contain signals for recycling or transcytosis meet in a common late endosomal compartment before being digested in lysosomes.

Summary

Cells ingest fluid, molecules, and particles by endocytosis, in which localized regions of the plasma membrane invaginate and pinch off to form endocytic vesicles. Many of the endocytosed molecules and particles eventually end up in lysosomes, where they are degraded. Endocytosis occurs both constitutively and as a triggered response to extracellular signals. Endocytosis is so extensive in many cells that a large fraction of the plasma membrane is internalized every hour. The cells remain the same size because most of the plasma membrane components (proteins and lipid) that are endocytosed are continually returned to the cell surface by exocytosis. This large-scale endocytic–exocytic cycle is mediated largely by clathrin-coated pits and vesicles.

Many cell-surface receptors that bind specific extracellular macromolecules become tagged with ubiquitin, which guides them into clathrin-coated pits. As a result, these receptors and their ligands are efficiently internalized in clathrin-coated vesicles, a process called receptor-mediated endocytosis. The coated vesicles rapidly shed their clathrin coats and fuse with early endosomes.

Most of the ligands dissociate from their receptors in the acidic environment of the endosome and eventually end up in lysosomes, while most of the receptors are recycled via transport vesicles back to the cell surface for reuse. But receptor–ligand complexes can follow other pathways from the endosomal compartment. In some cases, both the receptor and the ligand end up being degraded in lysosomes, resulting in receptor down-regulation; in these cases, the ubiquitin-tagged receptors recruit various ESCRT complexes, which drive the invagination and pinching-off of endosomal membrane vesicles to form multivesicular bodies. In other cases, both receptor and ligand are transferred to a different plasma membrane domain, causing the ligand to be released at a surface of the cell that differs from the membrane where it originated, a process called transcytosis. The transcytosis pathway involves recycling endosomes, where endocytosed plasma membrane proteins can be stored until they are needed.



TRANSPORT FROM THE *TRANS* GOLGI NETWORK TO THE CELL EXTERIOR: EXOCYTOSIS

Having considered the cell's internal digestive system and the various types of incoming membrane traffic that converge on lysosomes, we now return to the Golgi apparatus and examine the secretory pathways that lead out to the cell exterior. Transport vesicles destined for the plasma membrane normally leave

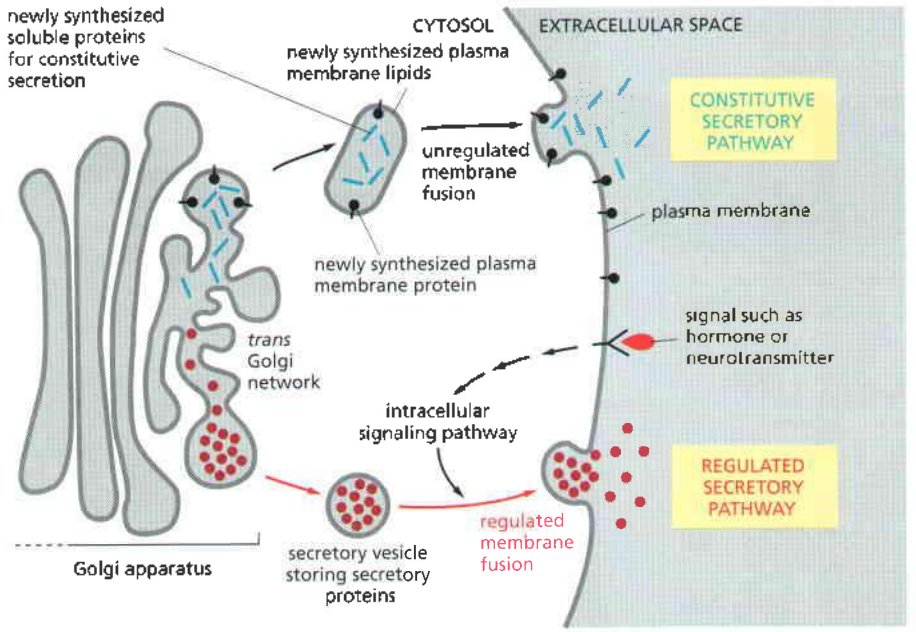


Figure 13–63 The constitutive and regulated secretory pathways. <ACAG> The two pathways diverge in the *trans* Golgi network. The constitutive secretory pathway operates in all cells. Many soluble proteins are continually secreted from the cell by this pathway, which also supplies the plasma membrane with newly synthesized lipids and proteins. Specialized secretory cells also have a regulated secretory pathway, by which selected proteins in the *trans* Golgi network are diverted into secretory vesicles, where the proteins are concentrated and stored until an extracellular signal stimulates their secretion. The regulated secretion of small molecules, such as histamine and neurotransmitters, occurs by a similar pathway; these molecules are actively transported from the cytosol into preformed secretory vesicles. There they are often complexed to specific macromolecules (proteoglycans, for histamine), so that they can be stored at high concentration without generating an excessively high osmotic pressure.

the TGN in a steady stream as irregularly shaped tubules. The membrane proteins and the lipids in these vesicles provide new components for the cell's plasma membrane, while the soluble proteins inside the vesicles are secreted to the extracellular space. The fusion of the vesicles with the plasma membrane is called **exocytosis.** In this way, for example, cells produce and secrete most of the proteoglycans and glycoproteins of the *extracellular matrix*, which is discussed in Chapter 19.

All cells require this **constitutive secretory pathway**, which operates continuously. Specialized secretory cells, however, have a second secretory pathway in which soluble proteins and other substances are initially stored in *secretory vesicles* for later release. This is the **regulated secretory pathway**, found mainly in cells specialized for secreting products rapidly on demand—such as hormones, neurotransmitters, or digestive enzymes (**Figure 13–63**). In this section, we consider the role of the Golgi apparatus in both of these secretory pathways and compare the two mechanisms of secretion.

Many Proteins and Lipids Seem to Be Carried Automatically from the Golgi Apparatus to the Cell Surface

A cell capable of regulated secretion must separate at least three classes of proteins before they leave the *trans* Golgi network—those destined for lysosomes (via endosomes), those destined for secretory vesicles, and those destined for immediate delivery to the cell surface (**Figure 13–64**). We have already noted that proteins destined for lysosomes are tagged for packaging into specific departing vesicles (with mannose 6-phosphate in the case of lysosomal hydrolases), and analogous signals are thought to direct *secretory proteins* into secretory vesicles. The non-selective constitutive secretory pathway transports most other proteins directly to the cell surface. Because entry into this pathway does not require a particular signal, it is also called the **default pathway.** Thus, in an unpolarized cell such as a white blood cell or a fibroblast, it seems that any protein in the lumen of the Golgi apparatus is automatically carried by the constitutive pathway to the cell surface unless it is specifically returned to the ER, retained as a resident protein in the Golgi apparatus itself, or selected for the pathways that lead to regulated secretion or to lysosomes. In polarized cells, where different products have to be delivered to different domains of the cell surface, we shall see that the options are more complex.

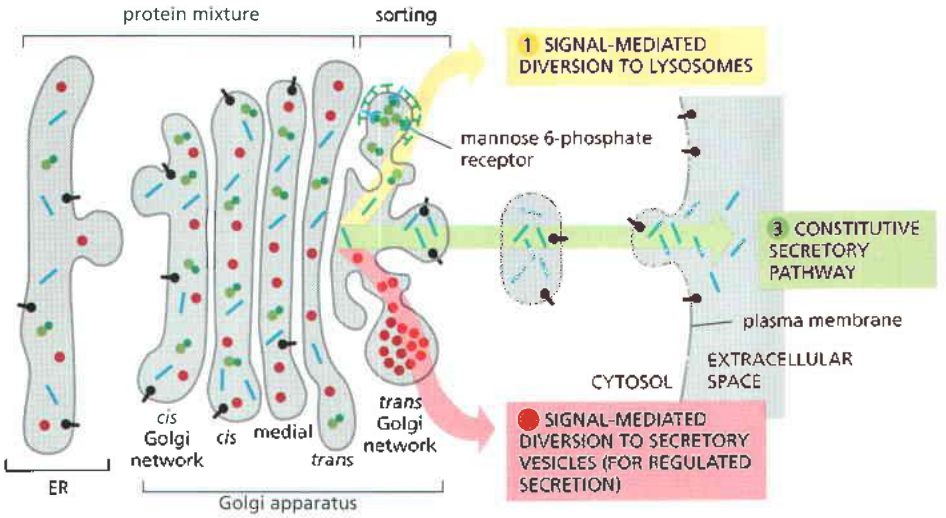


Figure 13–64 The three best-understood pathways of protein sorting in the *trans* Golgi network. (1) Proteins with the mannose 6-phosphate (M6P) marker are diverted to lysosomes (via endosomes) in clathrin-coated transport vesicles (see Figure 13–44). (2) Proteins with signals directing them to secretory vesicles are concentrated in such vesicles as part of a regulated secretory pathway that is present only in specialized secretory cells. (3) In unpolarized cells, a constitutive secretory pathway delivers proteins with no special features to the cell surface. In polarized cells, such as epithelial cells, however, secreted and plasma membrane proteins are selectively directed to either the apical or the basolateral plasma membrane domain, so a specific signal must mediate at least one of these two pathways, as we discuss later.

Secretory Vesicles Bud from the *Trans* Golgi Network

Cells that are specialized for secreting some of their products rapidly on demand concentrate and store these products in **secretory vesicles** (often called *secretory granules* or *dense-core vesicles* because they have dense cores when viewed in the electron microscope). Secretory vesicles form from the *trans* Golgi network, and they release their contents to the cell exterior by exocytosis in response to specific signals. The secreted product can be either a small molecule (such as histamine) or a protein (such as a hormone or digestive enzyme).

Proteins destined for secretory vesicles (called *secretory proteins*) are packaged into appropriate vesicles in the *trans* Golgi network by a mechanism that is thought to involve the selective aggregation of the secretory proteins. Clumps of aggregated, electron-dense material can be detected by electron microscopy in the lumen of the *trans* Golgi network. The signal that directs secretory proteins into such aggregates is not known, but it is thought to be composed of signal patches that are common to proteins of this class. When a gene encoding a secretory protein is artificially expressed in a secretory cell that normally does not make the protein, the foreign protein is appropriately packaged into secretory vesicles. This observation shows that, although the proteins that an individual cell expresses and packages in secretory vesicles differ, they all contain common sorting signals, which function properly even when the proteins are expressed in cells that do not normally make them.

It is unclear how the aggregates of secretory proteins are segregated into secretory vesicles. Secretory vesicles have unique proteins in their membrane, some of which might serve as receptors for aggregated protein in the *trans* Golgi network. The aggregates are much too big, however, for each molecule of the secreted protein to be bound by its own cargo receptor, as proposed for transport of the lysosomal enzymes. The uptake of the aggregates into secretory vesicles may therefore more closely resemble the uptake of particles by phagocytosis at the cell surface, where the plasma membrane zippers up around large structures.

Initially, most of the membrane of the secretory vesicles that leave the *trans* Golgi network is only loosely wrapped around the clusters of aggregated secretory proteins. Morphologically, these **immature secretory vesicles** resemble dilated *trans* Golgi cisternae that have pinched off from the Golgi stack. As the vesicles mature, they can fuse with one another and their contents become concentrated (**Figure 13–65**A), probably as the result of both the continuous retrieval of membrane that is recycled back to late endosomes and the TGN and the progressive acidification of the vesicle lumen that results from the increasing concentration of ATP-driven H^+ pumps in the vesicle membrane. Recall that V-type ATPases acidify all endocytic and exocytic organelles (see Figure 13–36). The degree of concentration of proteins during the formation and maturation

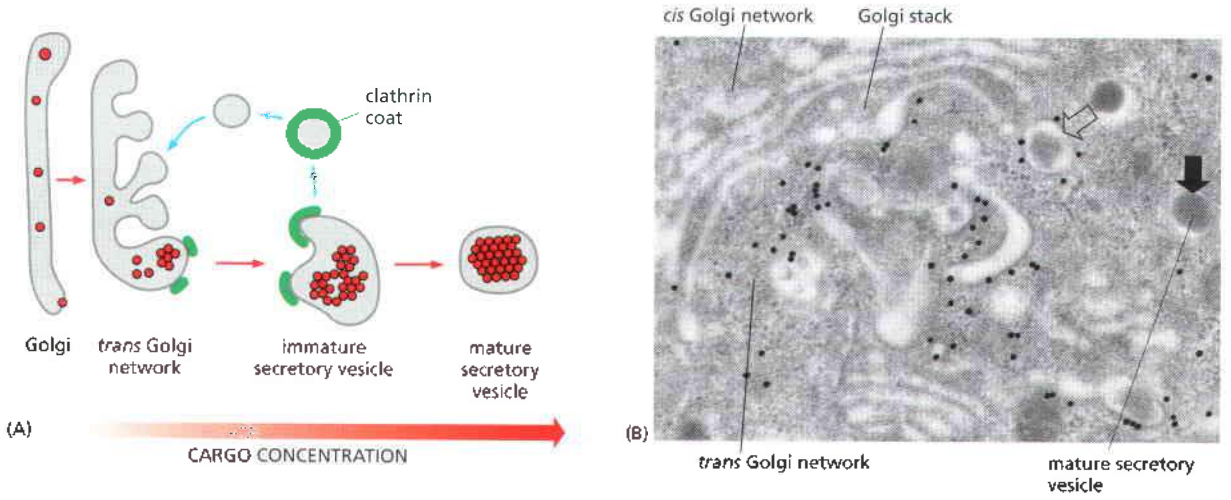


Figure 13–65 The formation of secretory vesicles. (A) Secretory proteins become segregated and highly concentrated in secretory vesicles by two mechanisms. First, they aggregate in the ionic environment of the *trans* Golgi network; often the aggregates become more condensed as secretory vesicles mature and their lumen becomes more acidic. Second, clathrin-coated vesicles retrieve excess membrane and lumenal content present in immature secretory vesicles as the secretory vesicles mature. (B) This electron micrograph shows secretory vesicles forming from the *trans* Golgi network in an insulin-secreting β cell of the pancreas. An antibody conjugated to gold spheres *(black dots)* has been used to locate clathrin molecules. The immature secretory vesicles *(open arrow)*, which contain insulin precursor protein (proinsulin), contain clathrin patches. Clathrin coats are no longer seen on the mature secretory vesicle, which has a highly condensed core *(solid arrow)*. (Courtesy of Lelio Orci.)

of secretory vesicles is only a small part of the total 200–400-fold concentration of these proteins that occurs after they leave the ER. Secretory and membrane proteins become concentrated as they move from the ER through the Golgi apparatus because of an extensive retrograde retrieval process mediated by COPI-coated transport vesicles that exclude them (see Figure 13–24).

Membrane recycling is important for returning Golgi components to the Golgi apparatus, as well as for concentrating the contents of secretory vesicles. The vesicles that mediate this retrieval originate as clathrin-coated buds on the surface of immature secretory vesicles, often being seen even on budding secretory vesicles that have not yet been severed from the Golgi stack (see Figure 13–65B).

Because the final mature secretory vesicles are so densely filled with contents, the secretory cell can disgorge large amounts of material promptly by exocytosis when triggered to do so (Figure 13–66).

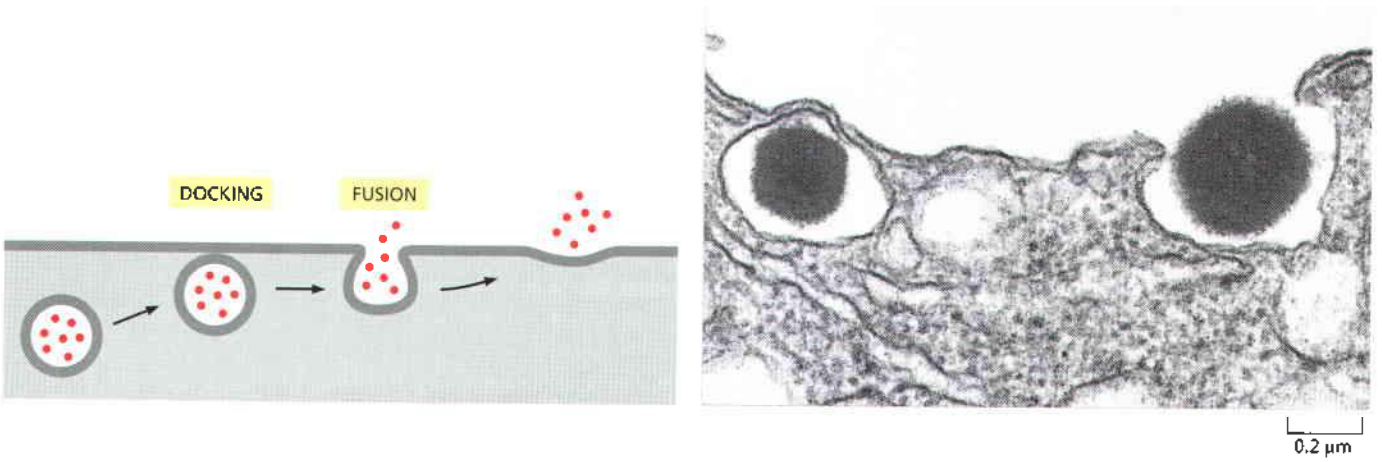


Figure 13–66 Exocytosis of secretory vesicles. The electron micrograph shows the release of insulin from a secretory vesicle of a pancreatic β cell. (Courtesy of Lelio Orci, from L. Orci, J.-D. Vassali and A. Perrelet, *Sci. Am.* 256:85–94, 1988. With permission from Scientific American.)

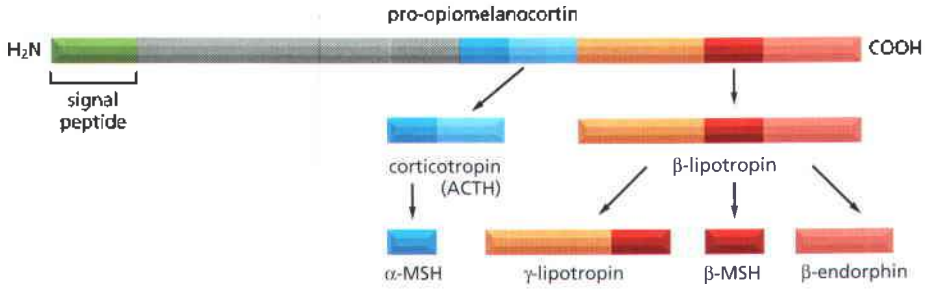


Figure 13–67 Alternative processing pathways for the prohormone pro-opiomelanocortin. The initial cleavages are made by proteases that cut next to pairs of positively charged amino acids (Lys-Arg, Lys-Lys, Arg-Lys, or Arg-Arg pairs). Trimming reactions then produce the final secreted products. Different cell types produce different concentrations of individual processing enzymes, so that the same prohormone precursor is cleaved to produce different peptide hormones. In the anterior lobe of the pituitary gland, for example, only corticotropin (ACTH) and β-lipotropin are produced from proopiomelanocortin, whereas in the intermediate lobe of the pituitary gland mainly α-melanocyte stimulating hormone (α-MSH), γ-lipotropin, β-MSH, and β-endorphin are produced.

Proteins Are Often Proteolytically Processed During the Formation of Secretory Vesicles

Concentration is not the only process to which secretory proteins are subject as the secretory vesicles mature. Many polypeptide hormones and neuropeptides, as well as many secreted hydrolytic enzymes, are synthesized as inactive protein precursors. Proteolysis is necessary to liberate the active molecules from these precursors. The cleavages begin in the *trans* Golgi network, and they continue in the secretory vesicles and sometimes in the extracellular fluid after secretion has occurred. Many secreted polypeptides have, for example, an N-terminal *propeptide* that is cleaved off to yield the mature protein. These proteins are thus synthesized as *pre-pro-proteins*, the *pre-peptide* consisting of the ER signal peptide that is cleaved off earlier in the rough ER (see Figure 12–38). In other cases, peptide-signaling molecules are made as *polyproteins* that contain multiple copies of the same amino acid sequence. In still more complex cases, a variety of peptide-signaling molecules are synthesized as parts of a single polyprotein that acts as a precursor for multiple end products, which are individually cleaved from the initial polypeptide chain. The same polyprotein may be processed in various ways to produce different peptides in different cell types (**Figure 13–67**).

Why is proteolytic processing so common in the secretory pathway? Some of the peptides produced in this way, such as the *enkephalins* (five-amino-acid neuropeptides with morphine-like activity), are undoubtedly too short in their mature forms to be co-translationally transported into the ER lumen or to include the necessary signal for packaging into secretory vesicles. In addition, for secreted hydrolytic enzymes—or any other protein whose activity could be harmful inside the cell that makes it—delaying activation of the protein until it reaches a secretory vesicle or until after it has been secreted has a clear advantage: it prevents it from acting prematurely inside the cell in which it is synthesized.

Secretory Vesicles Wait Near the Plasma Membrane Until Signaled to Release Their Contents

Once loaded, a secretory vesicle has to reach the site of secretion, which in some cells is far away from the Golgi apparatus. Nerve cells are the most extreme example. Secretory proteins, such as peptide neurotransmitters (neuropeptides) that are to be released from nerve terminals at the end of the axon, are made and packaged into vesicles in the cell body, where the ribosomes, ER, and Golgi apparatus are located. They must then travel along the axon to the nerve terminals, which can be a meter or more away. As discussed in Chapter 16, motor proteins propel the vesicles along axonal microtubules, whose uniform orientation guides the vesicles in the proper direction. Microtubules also guide vesicles to the cell surface for constitutive exocytosis.

Whereas vesicles containing materials for constitutive release fuse with the plasma membrane once they arrive there, secretory vesicles in the regulated pathway wait at the membrane until the cell receives a signal to secrete, and they then fuse. The signal is often a chemical messenger, such as a hormone, that binds to receptors on the cell surface. The resulting activation of the receptors

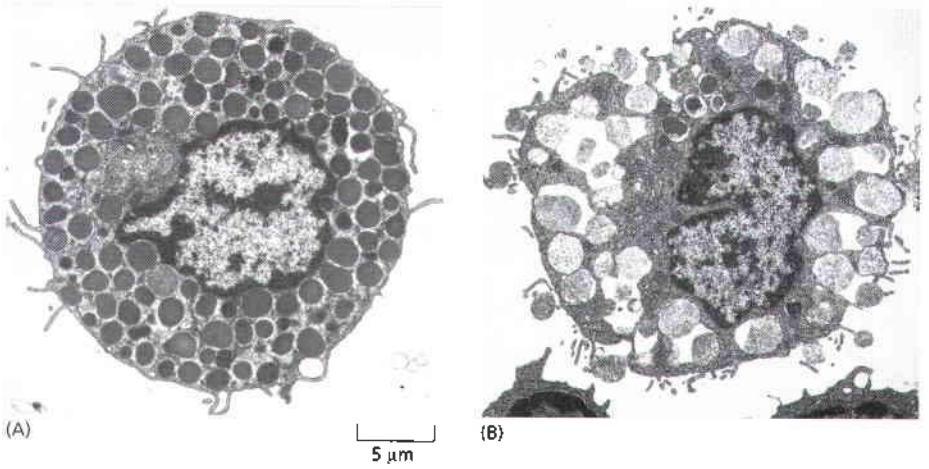


Figure 13–68 Electron micrographs of exocytosis in rat mast cells. (A) An unstimulated mast cell. (B) This cell has been activated to secrete its stored histamine by a soluble extracellular stimulant. Histamine-containing secretory vesicles are dark, while those that have released their histamine are light. The material remaining in the spent vesicles consists of a network of proteoglycans to which the stored histamine was bound. Once a secretory vesicle has fused with the plasma membrane, the secretory vesicle membrane often serves as a target to which other secretory vesicles fuse. Thus, the cell in (B) contains several large cavities lined by the fused membranes of many spent secretory vesicles, which are now in continuity with the plasma membrane. This continuity is not always apparent in one plane of section through the cell. (From D. Lawson, C. Fewtrell, B. Gomperts and M. Raff, *J. Exp. Med.* 142:391–402, 1975. With permission from The Rockefeller University Press.)

generates intracellular signals, often including a transient increase in the concentration of free Ca^{2+} in the cytosol. In nerve terminals, the initial signal for exocytosis is usually an electrical excitation (an action potential) triggered by a chemical transmitter binding to receptors elsewhere on the same cell surface. When the action potential reaches the nerve terminals, it causes an influx of Ca^{2+} through voltage-gated Ca^{2+} channels. The binding of Ca^{2+} ions to specific sensors then triggers the secretory vesicles (called synaptic vesicles) to fuse with the plasma membrane and release their contents to the extracellular space (see Figure 11–35).

The speed of transmitter release (taking only milliseconds) indicates that the proteins mediating the fusion reaction do not undergo complex, multistep rearrangements. After vesicles have been docked to the presynaptic plasma membrane, they undergo a priming step, which prepares them for rapid fusion. The SNAREs may be partly paired, but their helices are not fully wound into the final four-helix bundle required for fusion (see Figure 13–18). Other proteins are thought to keep the SNAREs from completing the fusion reaction until the Ca^{2+} influx releases this brake. At a typical synapse, only few of the docked vesicles seem to be primed and ready for exocytosis. The use of only a few vesicles at a time allows each synapse to fire over and over again in quick succession. With each firing, new synaptic vesicles become primed to replace those that have fused and released their contents.

Regulated Exocytosis Can Be a Localized Response of the Plasma Membrane and Its Underlying Cytoplasm

Histamine is a small molecule secreted by *mast cells*. It is released by the regulated pathway in response to specific ligands that bind to receptors on the mast cell surface (see Figure 25–27). Histamine causes many of the unpleasant symptoms that accompany allergic reactions, such as itching and sneezing. When mast cells are incubated in fluid containing a soluble stimulant, massive exocytosis occurs all over the cell surface (**Figure 13–68**). But if the stimulating ligand is artificially attached to a solid bead so that it can interact only with a localized region of the mast cell surface, exocytosis is now restricted to the region where the cell contacts the bead (**Figure 13–69**).

This experiment shows that individual segments of the plasma membrane can function independently in regulated exocytosis. As a result, the mast cell, unlike a nerve cell, does not respond as a whole when it is triggered; the activation of receptors, the resulting intracellular signals, and the subsequent exocytosis are all localized in the particular region of the cell that has been excited. Similarly, localized exocytosis enables a killer lymphocyte to deliver the proteins that induce the death of a single infected target cell precisely, without endangering normal neighboring cells (see Figure 25–46).

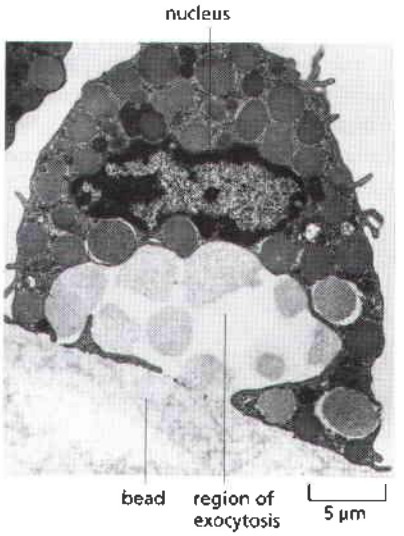


Figure 13–69 Exocytosis as a localized response. This electron micrograph shows a mast cell that has been activated to secrete histamine by a stimulant coupled to a large solid bead. Exocytosis has occurred only in the region of the cell that is in contact with the bead. (From D. Lawson, C. Fewtrell and M. Raff, *J. Cell Biol.* 79:394–400, 1978. With permission from The Rockefeller University Press.)

Secretory Vesicle Membrane Components Are Quickly Removed from the Plasma Membrane

When a secretory vesicle fuses with the plasma membrane, its contents are discharged from the cell by exocytosis, and its membrane becomes part of the plasma membrane. Although this should greatly increase the surface area of the plasma membrane, it does so only transiently, because membrane components are removed from the surface by endocytosis almost as fast as they are added by exocytosis, a process reminiscent of the exocytosis–endocytosis cycle discussed earlier. After their removal from the plasma membrane, the proteins of the secretory vesicle membrane are either recycled or shuttled to lysosomes for degradation. The amount of secretory vesicle membrane that is temporarily added to the plasma membrane can be enormous: in a pancreatic acinar cell discharging digestive enzymes for delivery to the gut lumen, about 900 μm^2 of vesicle membrane is inserted into the apical plasma membrane (whose area is only 30 μm^2) when the cell is stimulated to secrete.

Control of membrane traffic thus has a major role in maintaining the composition of the various membranes of the cell. To maintain each membrane-enclosed compartment in the secretory and endocytotic pathways at a constant size, the balance between the outward and inward flows of membrane needs to be precisely regulated. For cells to grow, the forward flow needs to be greater than the retrograde flow, so that the membrane can increase in area. For cells to maintain a constant size, the forward and retrograde flows must be equal. We still know very little about the mechanisms that coordinate these flows.

Some Regulated Exocytosis Events Serve to Enlarge the Plasma Membrane

An important task of regulated exocytosis is to deliver more membrane to enlarge the surface area of a cell's plasma membrane when such a need arises. A spectacular example is the plasma membrane expansion that occurs during the cellularization process of a fly embryo, which initially is a single cell containing about 6000 nuclei surrounded by a single plasma membrane. Within tens of minutes, the embryo is converted into the same number of cells. This process of cellularization requires a vast amount of new plasma membrane, which is added by a carefully orchestrated fusion of vesicles, eventually forming the plasma membranes that enclose the separate cells. Similar vesicle fusion events are required to enlarge the plasma membrane at the cleavage furrows of other animal and plant cells during *cytokinesis*, the process by which the two daughter cells separate after mitosis (discussed in Chapter 17).

Many cells, especially those subjected to mechanical stresses, frequently experience small ruptures to their plasma membranes during their life span. In a remarkable process thought to involve both homotypic vesicle–vesicle fusion and exocytosis, a temporary cell-surface patch is quickly fashioned from locally available internal-membrane sources, such as lysosomes. In addition to providing an emergency barrier against leaks, the patch also helps to reduce membrane tension over the wounded area allowing the "normal" bilayer to flow back together to restore continuity and seal the puncture. In wound repair, the fusion and exocytosis of vesicles is triggered by the sudden increase of Ca^{2+}, which is abundant in the extracellular space and rushes into the cells as soon as the plasma membrane is punctured. **Figure 13–70** shows three examples in which regulated exocytosis leads to plasma membrane expansion.

Polarized Cells Direct Proteins from the *Trans* Golgi Network to the Appropriate Domain of the Plasma Membrane

Most cells in tissues are *polarized* and have two (and sometimes more) distinct plasma membrane domains that are the targets of different types of vesicles.

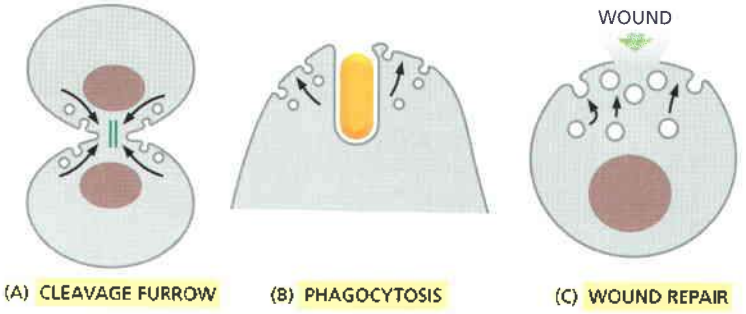


Figure 13–70 Three examples of regulated exocytosis leading to plasma membrane enlargement. (A, B) The vesicles fusing with the plasma membrane during cytokinesis and phagocytosis are thought to be derived from endosomes, whereas (C) those involved in wound repair may be derived from lysosomes. Plasma membrane wound repair is very important in cells subjected to mechanical stresses, such as muscle cells.

This raises the general problem of how the delivery of membrane from the Golgi apparatus is organized so as to maintain the differences between one cell-surface domain and another. A typical epithelial cell has an *apical domain*, which faces an internal cavity, or the outside world, and often has specialized features such as cilia or a brush border of microvilli; it also has a *basolateral domain*, which covers the rest of the cell. The two domains are separated by a ring of *tight junctions* (see Figure 19–24), which prevent proteins and lipids (in the outer leaflet of the lipid bilayer) from diffusing between the two domains, so that the compositions of the two domains are different.

A nerve cell is another example of a polarized cell. The plasma membrane of its axon and nerve terminals is specialized for signaling to other cells, whereas the plasma membrane of its cell body and dendrites is specialized for receiving signals from other nerve cells. The two domains have distinct protein compositions. Studies of protein traffic in nerve cells in culture suggest that, with regard to vesicular transport from the *trans* Golgi network to the cell surface, the plasma membrane of the nerve cell body and dendrites resembles the basolateral membrane of a polarized epithelial cell, while the plasma membrane of the axon and its nerve terminals resembles the apical membrane of such a cell (**Figure 13–71**). Thus, some proteins that are targeted to a specific domain in the epithelial cell are also targeted to the corresponding domain in the nerve cell.

Different Strategies Guide Membrane Proteins and Lipids Selectively to the Correct Plasma Membrane Domains

In principle, differences between plasma membrane domains need not depend on the targeted delivery of the appropriate membrane components. Instead, membrane components could be delivered to all regions of the cell surface indiscriminately but then be selectively stabilized in some locations and selectively eliminated in others. Although this strategy of random delivery followed by selective retention or removal seems to be used in certain cases, deliveries are often specifically directed to the appropriate membrane domain. Epithelial cells, for example, frequently secrete one set of products—such as digestive enzymes or mucus in cells lining the gut—at their apical surface and another set of products—such as components of the basal lamina—at their basolateral sur-

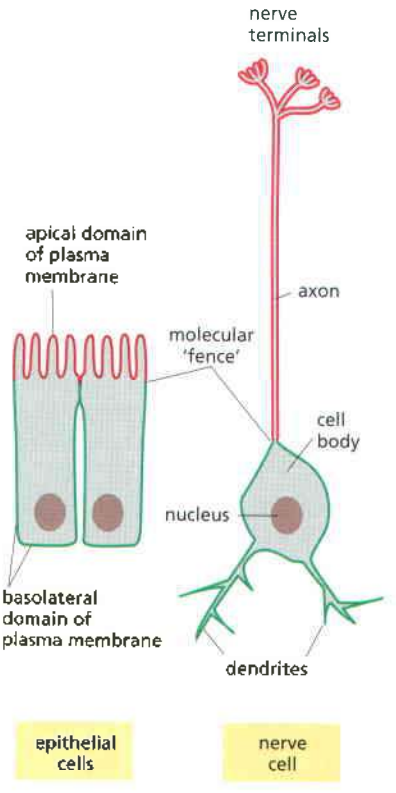


Figure 13–71 A comparison of two types of polarized cells. In terms of the mechanisms used to direct proteins to them, the plasma membrane of the nerve cell body and dendrites resembles the basolateral plasma membrane domain of a polarized epithelial cell, whereas the plasma membrane of the axon and its nerve terminals resembles the apical domain of an epithelial cell. The different membrane domains of both the epithelial cell and the nerve cell are separated by a molecular fence, consisting of a meshwork of membrane proteins tightly associated with the underlying actin cytoskeleton; this barrier—called a tight junction in the epithelial cell and an axonal hillock in neurons—keeps membrane proteins from diffusing between the two distinct domains.

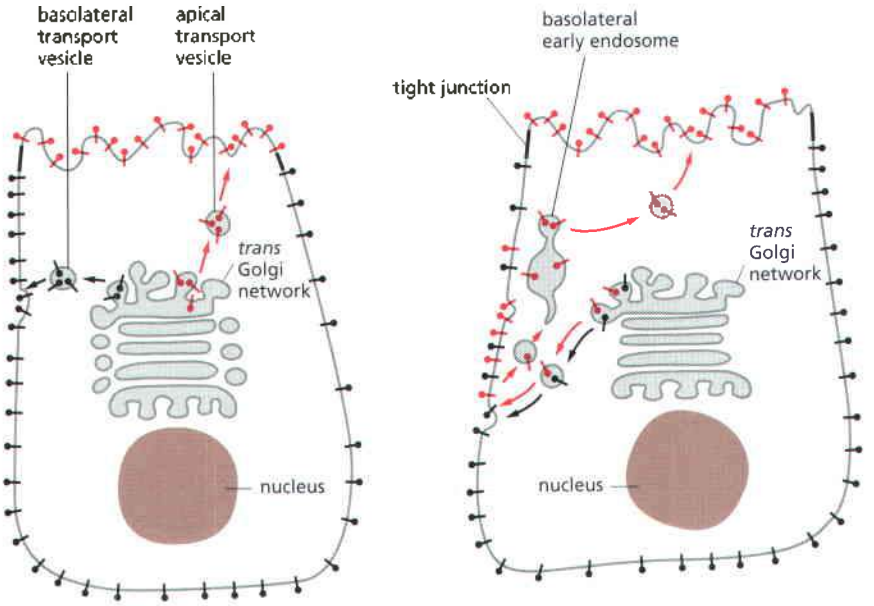


Figure 13–72 Two ways of sorting plasma membrane proteins in a polarized epithelial cell. Newly synthesized proteins can reach their proper plasma membrane domain by either (A) a direct specific pathway or (B) an indirect pathway. In the indirect pathway, a protein is retrieved from the inappropriate plasma membrane domain by endocytosis and then transported to the correct domain via early endosomes—that is, by transcytosis. The indirect pathway is used in liver hepatocytes to deliver proteins to the apical domain that lines bile ducts. However, in other cases, the direct pathway is used, as in the case of the lipid-raft-dependent mechanism in epithelial cells described in the text.

face. Thus, cells must have ways of directing vesicles carrying different cargoes to different plasma membrane domains.

By examining polarized epithelial cells in culture, it has been found that proteins from the ER destined for different domains travel together until they reach the TGN. Here they are separated and dispatched in secretory or transport vesicles to the appropriate plasma membrane domain (**Figure 13–72**).

The apical plasma membrane of most epithelial cells is greatly enriched in glycosphingolipids, which help protect this exposed surface from damage—for example, from the digestive enzymes and low pH in sites such as the gut or stomach, respectively. Similarly, plasma membrane proteins that are linked to the lipid bilayer by a glycosylphosphatidylinositol (GPI) anchor (discussed in Chapter 12) are found predominantly in the apical plasma membrane. If recombinant DNA techniques are used to attach a GPI anchor to a protein that would normally be delivered to the basolateral surface, the protein is usually delivered to the apical surface instead. GPI-anchored proteins are thought to be directed to the apical membrane because they associate with glycosphingolipids in lipid rafts that form in the membrane of the TGN. As discussed in Chapter 10, lipid rafts form in the TGN and plasma membrane when glycosphingolipids and cholesterol self-associate into microaggregates (see Figure 10–14). Having selected a unique set of cargo molecules, the rafts then bud from the *trans* Golgi network into transport vesicles destined for the apical plasma membrane. Thus, similar to the selective partitioning of some membrane proteins into the specialized lipid domains in caveolae at the plasma membrane, lipid domains may also participate in protein sorting in the TGN.

Membrane proteins destined for delivery to the basolateral membrane contain sorting signals in their cytosolic tail. When present in an appropriate structural context, these signals are recognized by coat proteins that package them into appropriate transport vesicles in the TGN. The same basolateral signals that are recognized in the TGN also function in endosomes to redirect the proteins back to the basolateral plasma membrane after they have been endocytosed.

Synaptic Vesicles Can Form Directly from Endocytic Vesicles

Nerve cells (and some endocrine cells) contain two types of secretory vesicles. As for all secretory cells, these cells package proteins and peptides in dense-cored

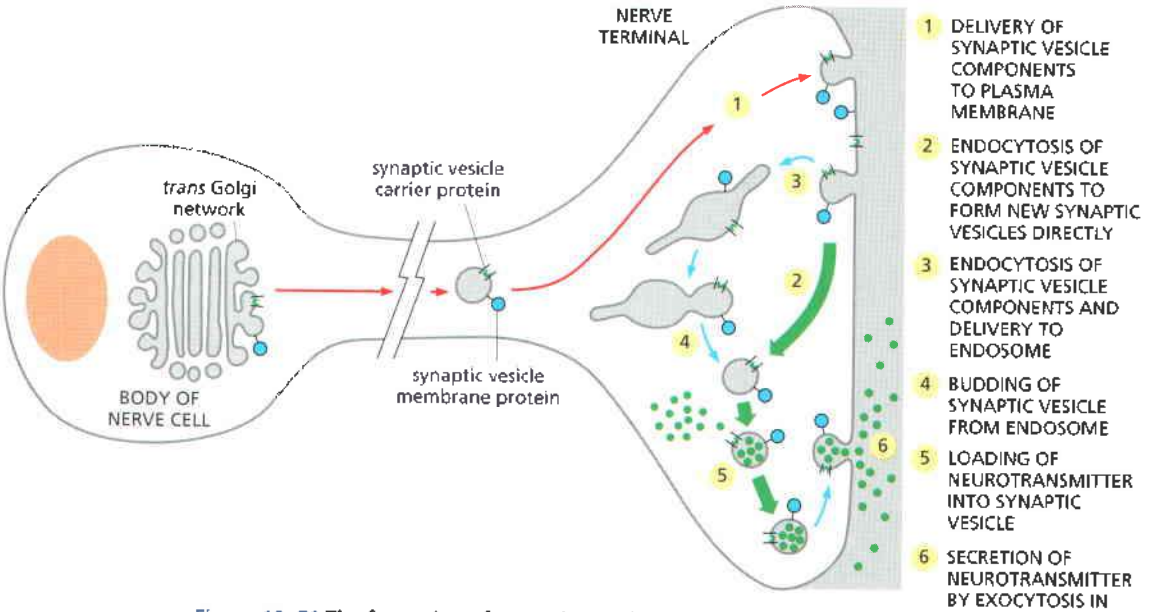


Figure 13–73 The formation of synaptic vesicles. These tiny uniform vesicles are found only in nerve cells and in some endocrine cells, where they store and secrete small-molecule neurotransmitters. The import of neurotransmitter directly into the small endocytic vesicles that form from the plasma membrane is mediated by membrane carrier proteins that function as antiports, being driven by a H^+ gradient maintained by proton pumps in the vesicle membrane.

secretory vesicles in the standard way for release by the regulated secretory pathway. In addition, however, they use another specialized class of tiny (~50-nm diameter) secretory vesicles called **synaptic vesicles**, which are generated in a different way. In nerve cells, these vesicles store small neurotransmitter molecules, such as acetylcholine, glutamate, glycine, and γ-aminobutyric acid (GABA), which mediate rapid signaling from cell to cell at chemical synapses. As discussed earlier, when an action potential arrives at a nerve terminal, it triggers the vesicles to release their contents within a fraction of a millisecond. Some neurons fire more than 1000 times per second, releasing neurotransmitters each time. This rapid release is possible because some of the vesicles are docked and primed for fusion, which will occur only when an action potential causes an influx of Ca^{2+} into the terminal.

Only a small proportion of the synaptic vesicles in the nerve terminal fuse with the plasma membrane in response to each action potential. But for the nerve terminal to respond rapidly and repeatedly, the vesicles need to be replenished very quickly after they discharge. Thus, most synaptic vesicles are generated not from the Golgi membrane in the nerve cell body but by local recycling from the plasma membrane in the nerve terminals. It is thought that the membrane components of the synaptic vesicles are initially delivered to the plasma membrane by the constitutive secretory pathway and then retrieved by endocytosis. But instead of fusing with endosomes, most of the endocytic vesicles immediately fill with transmitter to become synaptic vesicles.

The membrane components of a synaptic vesicle include transporters specialized for the uptake of neurotransmitter from the cytosol, where the small-molecule neurotransmitters that mediate fast synaptic signaling are synthesized. Once filled with neurotransmitter, the vesicles return to the plasma membrane, where they wait until the cell is stimulated. After they have released their contents, their membrane components are retrieved in the same way and used again (Figure 13–73).

Because synaptic vesicles are abundant and relatively uniform in size, they can be purified in large numbers and, consequently, are the best-characterized organelle of the cell. Careful quantitative proteomic analyses have identified all the components of a synaptic vesicle (**Figure 13–74**).

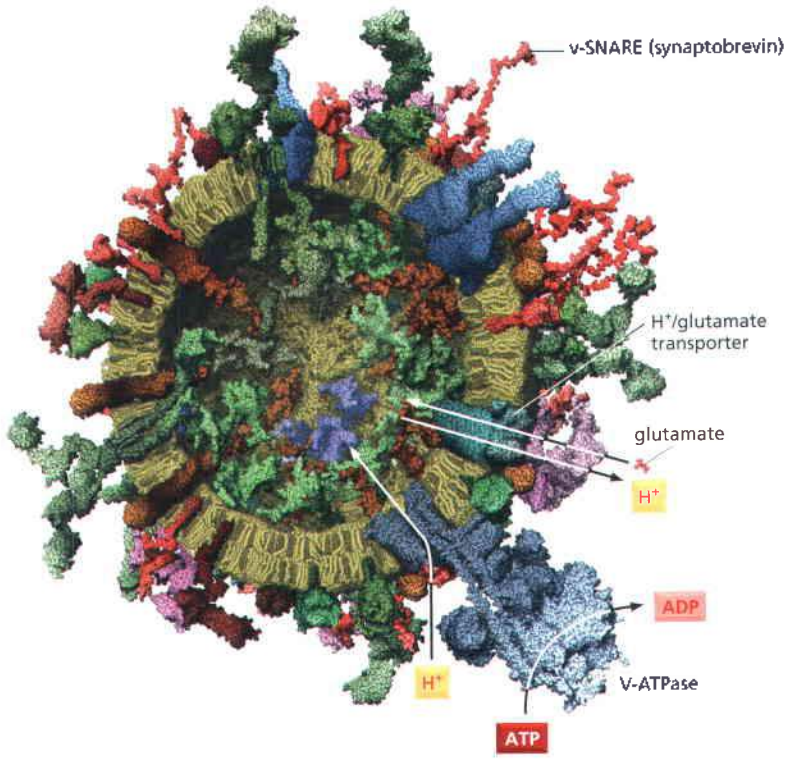


Figure 13–74 A scale model of a synaptic vesicle. The illustration shows a section through a synaptic vesicle in which proteins and lipids are drawn to scale based on their known stoichiometry and either their known or approximated structures. Only 70% of the membrane proteins estimated to be present in the membrane are shown; a complete model would therefore show a membrane that is even more crowded than this picture suggests. Each synaptic vesicle membrane contains 7000 phospholipid molecules and 5700 cholesterol molecules. Each also contains close to 50 different integral membrane protein molecules, which vary widely in their relative abundance and together contribute about 600 transmembrane α-helices. The v-SNARE synaptobrevin is the most abundant protein in the vesicle (~70 copies/vesicle). Because of its filamentous structure, it extends from the dense forest of the cytosolic protein domains, which almost completely cover the vesicle surface. By contrast, the V-ATPase, which uses ATP hydrolysis to pump H^+ into the vesicle lumen, is present in 1–2 copies per vesicle. The H^+ gradient provides the energy for neurotransmitter import by an H^+/neurotransmitter antiport, which loads each vesicle with 1800 neurotransmitter molecules, such as glutamate, one of which is shown to scale. (Adapted from S. Takamori et al., *Cell*, 127:831–846, 2006. With permission from Elsevier.) <ACCA>

Summary

Cells can secrete molecules by exocytosis in either a constitutive or a regulated fashion. Whereas the regulated pathways operate only in specialized secretory cells, a constitutive secretory pathway operates in all eucaryotic cells, characterized by continual vesicular transport from the TGN to the plasma membrane. In the regulated pathways, the molecules are stored either in secretory vesicles or in synaptic vesicles, which do not fuse with the plasma membrane to release their contents until they receive an appropriate signal. Secretory vesicles containing proteins for secretion bud from the TGN. The secretory proteins become concentrated during the formation and maturation of the secretory vesicles. Synaptic vesicles, which are confined to nerve cells and some endocrine cells, form from both endocytic vesicles and from endosomes, and they mediate the regulated secretion of small-molecule neurotransmitters.

Proteins are delivered from the TGN to the plasma membrane by the constitutive pathway unless they are diverted into other pathways or retained in the Golgi apparatus. In polarized cells, the transport pathways from the TGN to the plasma membrane operate selectively to ensure that different sets of membrane proteins, secreted proteins, and lipids are delivered to the different domains of the plasma membrane.

PROBLEMS

Which statements are true? Explain why or why not.

13–1 In all events involving fusion of a vesicle to a target membrane, the cytosolic leaflets of the vesicle and target bilayers always fuse together, as do the leaflets that are not in contact with the cytosol.

13–2 There is one strict requirement for the exit of a protein from the ER: it must be correctly folded.

13–3 All of the glycoproteins and glycolipids in intracellular membranes have their oligosaccharide chains facing the lumenal side, and all those in the plasma membrane have their oligosaccharide chains facing the outside of the cell.

13–4 During transcytosis, vesicles that form from coated pits on the apical surface fuse with the plasma membrane on the basolateral surface, and in that way transport molecules across the epithelium.

Discuss the following problems.

13–5 In a nondividing cell such as a liver cell, why must the flow of membrane between compartments be balanced, so that the retrieval pathways match the outward flow? Would you expect the same balanced flow in a gut epithelial cell, which is actively dividing?

13–6 Yeast, and many other organisms, make a single type of clathrin heavy chain and a single type of clathrin light chain; thus, they make a single kind of clathrin coat. How is it, then, that a single clathrin coat can be used for three different transport pathways—Golgi to late endosomes, plasma membrane to early endosomes, and immature secretory vesicles to Golgi—that each involves different specialized cargo proteins?

13–7 How can it possibly be true that complementary pairs of specific SNAREs uniquely mark vesicles and their target membranes? After vesicle fusion, the target membrane will contain a mixture of t-SNAREs and v-SNAREs. Initially, these SNAREs will be tightly bound to one another, but NSF can pry them apart, reactivating them. What do you suppose prevents target membranes from accumulating a population of v-SNAREs equal to or greater than their population of t-SNAREs?

13–8 Viruses are the ultimate scavengers—a necessary consequence of their small genomes. Wherever possible, they make use of the cell's machinery to accomplish the steps involved in their own reproduction. Many different viruses have membrane coverings. These so-called enveloped viruses gain access to the cytosol by fusing with a cell membrane. Why do you suppose that each of these viruses encodes its own special fusion protein, rather than making use of a cell's SNAREs?

13–9 For fusion of a vesicle with its target membrane to occur, the membranes have to be brought to within 1.5 nm so that the two bilayers can join (**Figure Q13–1**). Assuming that the relevant portions of the two membranes at the fusion site are circular regions 1.5 nm in diameter, calculate the number of water molecules that would remain between the membranes. (Water is 55.5 M and the volume of a cylinder is $\pi r^2 h$.) Given that an average phospholipid occupies a membrane surface area of 0.2 nm^2, how many phospholipids would be present in each of the opposing monolayers at the fusion site? Are there sufficient water molecules to bind to the hydrophilic head groups of this number of phospholipids? (It is estimated that 10–12 water molecules are normally associated with each phospholipid head group at the exposed surface of a membrane.)

13–10 SNAREs exist as complementary partners that carry out membrane fusions between appropriate vesicles and their target membranes. In this way, a vesicle with a particular variety of v-SNARE will fuse only with a membrane that carries the complementary t-SNARE. In some instances, however, fusions of identical membranes (homotypic fusions) are known to occur. For example, when a yeast cell forms a bud, vesicles derived from the mother cell's vacuole move into the bud where they fuse with one another to form a new vacuole. These vesicles carry both v-SNAREs and t-SNAREs. Are both types of SNAREs essential for this homotypic fusion event?

To test this point, you have developed an ingenious assay for fusion of vacuolar vesicles. You prepare vesicles from two different mutant strains of yeast: strain B has a defective gene for vacuolar alkaline phosphatase (Pase); strain A is defective for the protease that converts the precursor of alkaline phosphatase (pro-Pase) into its active form (Pase) (**Figure Q13–2A**). Neither strain has active alkaline phosphatase, but when extracts of the strains are mixed, vesicle fusion generates active alkaline phosphatase, which can be easily measured (Figure Q13–2).

Now you delete the genes for the vacuolar v-SNARE, t-SNARE, or both in each of the two yeast strains. You prepare vacuolar vesicles from each and test them for their ability to fuse, as measured by the alkaline phosphatase assay (Figure Q13–2B).

What do these data say about the requirements for v-SNAREs and t-SNAREs in the fusion of vacuolar vesicles? Does it matter which kind of SNARE is on which vesicle?

13–11 If you were to remove the ER retrieval signal from protein disulfide isomerase (PDI), which is normally a soluble resident of the ER lumen, where would you expect the modified PDI to be located?

13–12 The KDEL receptor must shuttle back and forth between the ER and the Golgi apparatus to accomplish its task of ensuring that soluble ER proteins are retained in the ER lumen. In which compartment does the KDEL receptor bind its ligands more tightly? In which compartment does it bind its ligands more weakly? What is thought to be the basis for its different binding affinities in the two compartments? If you were designing the system, in which compartment would you have the highest concentration of KDEL receptor? Would you predict that the KDEL receptor,

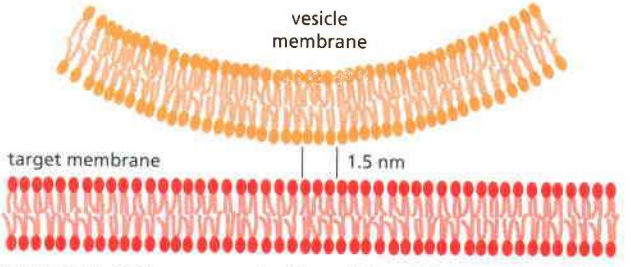


Figure Q13–1 Close approach of a vesicle and its target membrane in preparation for fusion (Problem 13–9).

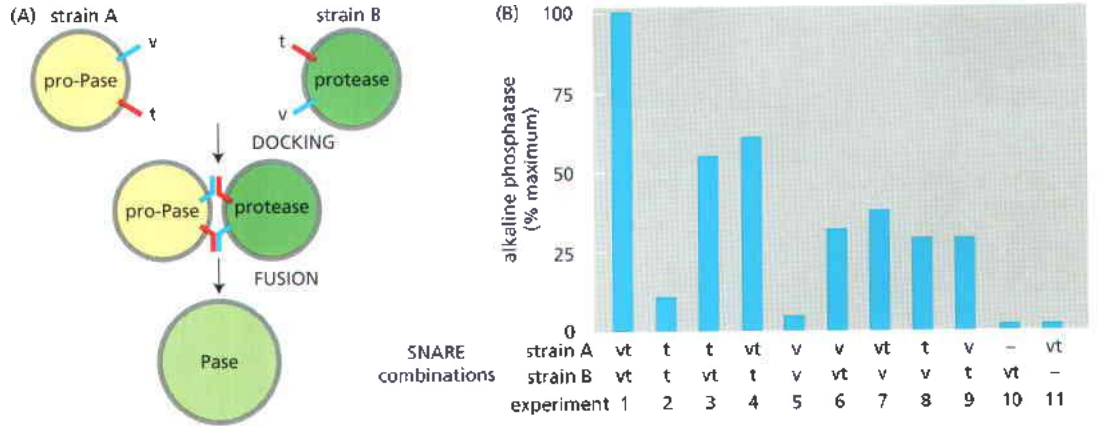


Figure Q13–2 SNARE requirements for vesicle fusion (Problem 13–10). (A) Scheme for measuring the fusion of vacuolar vesicles. (B) Results of fusions of vesicles with different combinations of v-SNAREs and t-SNAREs. The SNAREs present on the vesicles of the two strains are indicated as v (v-SNARE) and t (t-SNARE).

which is a transmembrane protein, would itself possess an ER retrieval signal?

13–13 How does the low pH of lysosomes protect the rest of the cell from lysosomal enzymes in case the lysosome breaks?

13–14 Melanosomes are specialized lysosomes that store pigments for eventual release by exocytosis. Various cells such as skin and hair cells then take up the pigment, which accounts for their characteristic pigmentation. Mouse mutants that have defective melanosomes often have pale or unusual coat colors. One such light-colored mouse, the *Mocha* mouse (**Figure Q13–3**), has a defect in the gene for one of the subunits of the adaptor protein complex AP3, which is associated with coated vesicles budding from the *trans* Golgi network. How might the loss of AP3 cause a defect in melanosomes?



Figure Q13–3 A normal mouse and the *Mocha* mouse (Problem 13–14). In addition to its light coat color, the *Mocha* mouse has a poor sense of balance. (Courtesy of Margit Burmeister.)

13–15 Patients with Hunter's syndrome or with Hurler's syndrome rarely live beyond their teens. These patients accumulate glycosaminoglycans in lysosomes due to the lack of specific lysosomal enzymes necessary for their degradation. When cells from patients with the two syndromes are fused, glycosaminoglycans are degraded properly, indicating that the cells are missing different degradative enzymes. Even if the cells are just cultured together, they still correct each other's defects. Most surprising of all, the medium from a culture of Hurler's cells corrects the defect in Hunter's cells (and vice versa). The corrective factors in the media are inactivated by treatment with proteases, by treatment with periodate, which destroys carbohydrate, and by treatment with alkaline phosphatase, which removes phosphates.

A. What do you suppose the corrective factors are? Beginning with the donor patient's cells, describe the route by which the factors reach the medium and subsequently enter the recipient cells to correct the lysosomal defects.

B. Why do you suppose the treatments with protease, periodate, and alkaline phosphatase inactivate the corrective factors?

C. Would you expect a similar sort of correction scheme to work for mutant cytosolic enzymes?

13–16 A macrophage ingests the equivalent of 100% of its plasma membrane each half hour by endocytosis. What is the rate at which membrane is returned by exocytosis?

13–17 Cells take up extracellular molecules by receptor-mediated endocytosis and by fluid-phase endocytosis. A classic paper compared the efficiencies of these two pathways by incubating human cells for various periods of time in a range of concentrations of either ^{125}I-labeled epidermal growth factor (EGF), to measure receptor-mediated endocytosis, or horseradish peroxidase (HRP), to measure fluid-phase endocytosis. Both EGF and HRP were found to be present in small vesicles with an internal radius of 20 nm. The uptake of HRP was linear (**Figure Q13–4A**), while that of EGF was initially linear but reached a plateau at higher concentrations (Figure Q13–4B).

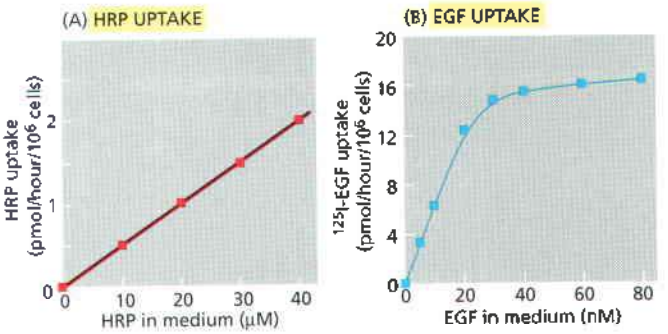


Figure Q13–4 Uptake of HRP and EGF as a function of their concentration in the medium (Problem 13–17).

A. Explain why the shapes of the curves in Figure Q13–4 are different for HRP and EGF.

B. From the curves in Figure Q13–4, estimate the difference in the uptake rates for HRP and EGF when both are present at 40 nM. What would the difference be if both were present at 40 μM?

C. Calculate the average number of HRP molecules that get taken up by each endocytic vesicle (radius 20 nm) when the medium contains 40 μM HRP. [The volume of a sphere is $(4/3)\pi r^3$.]

D. The scientists who did these experiments said at the time, "These calculations clearly illustrate how cells can internalize EGF by endocytosis while excluding all but insignificant quantities of extracellular fluid." What do you think they meant.

REFERENCES

General

Bonifacino JS & Glick BS (2004) The mechanisms of vesicle budding and fusion. *Cell* 116:153–66.

Mellman I & Warren G (2000) The road taken: past and future foundations of membrane traffic. *Cell* 100:99–112.

Pelham HR (1999) The Croonian Lecture 1999. Intracellular membrane traffic: getting proteins sorted. *Philos Trans R Soc Lond B Biol Sci* 354:1471–1478.

Rothman JE & Wieland FT (1996) Protein sorting by transport vesicles. *Science* 272:227–234.

Schekman RW (1994) Regulation of membrane traffic in the secretory pathway. *Harvey Lect* 90:41–57.

The Molecular Mechanisms of Membrane Transport and the Maintenance of Compartment Diversity

Di Paolo G & De Camilli P (2006) Phosphoinositides in cell regulation and membrane dynamics. *Nature* 443:651–657.

Grosshans BL, Ortiz D & Novick P (2006) Rabs and their effectors: achieving specificity in membrane traffic. *Proc Natl Acad Sci USA* 103:11821–11827.

Gurkan C, Stagg SM & Balch WE (2006) The COPII cage: unifying principles of vesicle coat assembly. *Nature Rev Mol Cell Biol* 7:727–738.

Jahn R & Scheller RH (2006) SNAREs—engines for membrane fusion. *Nature Rev Mol Cell Biol* 7:631–643.

Kirchhausen T (2000) Three ways to make a vesicle. *Nature Rev Mol Cell Biol* 1:187–198.

McNew JA, Parlati F, Sollner T & Rothman JE (2000) Compartmental specificity of cellular membrane fusion encoded in SNARE proteins. *Nature* 407:153–159.

Pfeffer SR (1999) Transport-vesicle targeting: tethers before SNAREs. *Nature Cell Biol* 1:E17–E22.

Robinson MS & Bonifacino JS (2001) Adaptor-related proteins. *Curr Opin Cell Biol* 13:444–453.

Seaman MN (2005) Recycle your receptors with retromer. *Trends Cell Biol* 15:68–75.

Springer S, Spang A & Schekman RW (1999) A primer on vesicle budding. *Cell* 97:145–148.

Transport from the ER Through the Golgi Apparatus

Bannykh SI, Nishimura N & Balch WE (1998) Getting into the Golgi. *Trends Cell Biol* 8:21–25.

Ellgaard L & Helenius A (2003) Quality control in the endoplasmic reticulum. *Nature Rev Mol Cell Biol* 4:181–191.

Farquhar MG & Palade GE (1998) The Golgi apparatus: 100 years of progress and controversy. *Trends Cell Biol* 8:2–10.

Glick BS (2000) Organization of the Golgi apparatus. *Curr Opin Cell Biol* 12:450–456.

Ladinsky MS, Mastronarde DN et al (1999) Golgi structure in three dimensions: functional insights from the normal rat kidney cell. *J Cell Biol* 144:1135–1149.

Lee MC, Miller EA, Goldberg J et al (2004) Bi-directional protein transport between the ER and Golgi. *Annu Rev Cell Dev Biol* 20:87–123.

Pelham HR & Rothman JE (2000) The debate about transport in the Golgi—two sides of the same coin? *Cell* 102:713–719.

Warren G & Malhotra V (1998) The organisation of the Golgi apparatus. *Curr Opin Cell Biol* 10:493–498.

Wells L & Hart GW (2003) O-GlcNAc turns twenty: functional implications for post-translational modification of nuclear and cytosolic proteins with a sugar. *FEBS Lett* 546:154–158.

Zeuschner D, Geerts WJ & Klumperman J (2006) Immuno-electron tomography of ER exit sites reveals the existence of free COPII-coated transport carriers. *Nature Cell Biol* 8:377–383.

Transport from the *Trans* Golgi Network to Lysosomes

Andrews NW (2000) Regulated secretion of conventional lysosomes. *Trends Cell Biol* 10:316–321.

de Duve C (2005) The lysosome turns fifty. *Nature Cell Biol* 7:847–849.

Kornfeld S & Mellman I (1989) The biogenesis of lysosomes. *Annu Rev Cell Biol* 5:483–525.

Levine B & Klionsky DJ (2004) Development by self-digestion: molecular mechanisms and biological functions of autophagy. *Dev Cell* 6:463–477.

Noda T, Suzuki K & Ohsumi Y (2002) Yeast autophagosomes: *de novo* formation of a membrane structure. *Trends Cell Biol* 12:231–235.

Futerman AH & van Meer G (2004) The cell biology of lysosomal storage disorders. *Nature Rev Mol Cell Biol* 5:554–565.

Peters C & von Figura K (1994) Biogenesis of lysosomal membranes. *FEBS Lett* 346:108–114.

Rouille Y, Rohn W & Hoflack B (2000) Targeting of lysosomal proteins. *Semin Cell Dev Biol* 11:165–171.

Transport into the Cell from the Plasma Membrane: Endocytosis

Anderson RG (1998) The caveolae membrane system. *Annu Rev Biochem* 67:199–225.

Bonifacino JS & Traub LM (2003) Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annu Rev Biochem* 72:395–447.

Brown MS & Goldstein JL (1986) A receptor-mediated pathway for cholesterol homeostasis. *Science* 232:34–47.

Conner SD & Schmid SL (2003) Regulated portals of entry into the cell. *Nature* 422:37–44.

Gruenberg J & Stenmark H (2004) The biogenesis of multivesicular endosomes. *Nature Rev Mol Cell Biol* 5:317–323.

Hicke L (2001) A new ticket for entry into budding vesicles-ubiquitin. *Cell* 106:527–530.

Katzmann DJ, Odorizzi G & Emr S (2002) Receptor downregulation and multivesicular-body sorting. *Nature Rev Mol Cell Biol* 3:893–905.

Maxfield FR & McGraw TE (2004) Endocytic recycling. *Nature Rev Mol Cell Biol* 5:121–132.

Mellman I (1996) Endocytosis and molecular sorting. *Annu Rev Cell Dev Biol* 12:575–625.

Pelkmans L & Helenius A (2003) Insider information: what viruses tell us about endocytosis. *Curr Opin Cell Biol* 15:414–422.

Tjelle TE, Lovdal T et al (2000) Phagosome dynamics and function. *BioEssays* 22:255–263.

Yeung T, Ozdamar B & Grinstein S (2006) Lipid metabolism and dynamics during phagocytosis. *Curr Opin Cell Biol* 18:429–437.

Transport from the *Trans* Golgi Network to the Cell Exterior: Exocytosis

Burgess TL & Kelly RB (1987) Constitutive and regulated secretion of proteins. *Annu Rev Cell Biol* 3:243–293.

Dietrich C, Volovyk ZN & Jacobson K (2001) Partitioning of Thy-1, GM1, and cross-linked phospholipid analogs into lipid rafts reconstituted in supported model membrane monolayers. *Proc Natl Acad Sci USA* 98:10642–10647.

Martin TF (1997) Stages of regulated exocytosis. *Trends Cell Biol* 7:271–276.

Mostov K, Su T et al (2003) Polarized epithelial membrane traffic: conservation and plasticity. *Nature Cell Biol* 5:287–293.

Murthy VN & De Camilli P (2003) Cell biology of the presynaptic terminal. *Annu Rev Neurosci* 26:701–728.

Schuck S & Simons K (2004) Polarized sorting in epithelial cells: raft clustering and the biogenesis of the apical membrane. *J Cell Sci* 117(Pt 25):5955–5964.

Simons K & Ikonen E (1997) Functional rafts in cell membranes. *Nature* 387:569–572.

Sudhof TC (2004) The synaptic vesicle cycle. *Annu Rev Neurosci* 27:509–547.

Tooze SA (1998) Biogenesis of secretory granules in the trans-Golgi network of neuroendocrine and endocrine cells. *Biochim Biophys Acta* 1404:231–244.

Traub LM & Kornfeld S (1997) The trans-Golgi network: a late secretory sorting station. *Curr Opin Cell Biol* 9:527–533.

Energy Conversion: Mitochondria and Chloroplasts

14

In This Chapter

THE MITOCHONDRION 815

ELECTRON-TRANSPORT CHAINS AND THEIR PROTON PUMPS 827

CHLOROPLASTS AND PHOTOSYNTHESIS 841

THE GENETIC SYSTEMS OF MITOCHONDRIA AND PLASTIDS 855

THE EVOLUTION OF ELECTRON-TRANSPORT CHAINS 870

Through a set of reactions that occur in the cytosol, energy derived from the partial oxidation of energy-rich carbohydrate molecules is used to form ATP, the chemical energy currency of cells (discussed in Chapter 2). But a much more efficient method of energy generation appeared very early in the history of life. This process requires a membrane, and it enables cells to acquire energy from a wide variety of sources. For example, it is central to the conversion of light energy into chemical bond energy in photosynthesis, as well as to the aerobic respiration that enables us to use oxygen to produce large amounts of ATP from food molecules.

Procaryotes use their plasma membrane to produce ATP. But the plasma membrane in eucaryotic cells is reserved for the transport processes described in Chapter 11. Eucaryotes instead use the specialized membranes inside *energy-converting organelles* to produce most of their ATP. The membrane-enclosed organelles are **mitochondria**, which are present in the cells of virtually all eucaryotic organisms (including fungi, animals, plants, algae, and protozoa), and **plastids**—most notably **chloroplasts**—which occur only in plants and algae. In electron micrographs the most striking morphological feature of both mitochondria and chloroplasts is the large amount of internal membrane they contain. This internal membrane provides the framework for an elaborate set of electron-transport processes that produce most of the cell's ATP.

The common pathway used by mitochondria, chloroplasts, and procaryotes to harness energy for biological purposes operates by a process known as **chemiosmotic coupling**—reflecting a link between the chemical bond-forming reactions that generate ATP ("chemi") and membrane-transport processes ("osmotic"). The coupling process occurs in two linked stages, both of which are performed by protein complexes embedded in a membrane:

Stage 1. High-energy electrons (derived from the oxidation of food molecules, from the action of sunlight, or from other sources discussed later) are transferred along a series of electron carriers embedded in the membrane. These electron transfers release energy that is used to pump protons (H^+, derived from the water that is plentiful in cells) across the membrane and thus generate an *electrochemical proton gradient*. As discussed in Chapter 11, an ion gradient across a membrane is a form of stored energy, which can be harnessed to do useful work when the ions are allowed to flow back across the membrane down their electrochemical gradient.

Stage 2. H^+ flows back down its electrochemical gradient through a protein machine called *ATP synthase*, which catalyzes the energy-requiring synthesis of ATP from ADP and inorganic phosphate (P_i). This ubiquitous enzyme plays the role of a turbine, permitting the proton gradient to drive the production of ATP (**Figure 14–1**).

The electrochemical proton gradient also drives other membrane-embedded protein machines (**Figure 14–2**). In eucaryotes, special proteins couple the "downhill" H^+ flow to the transport of specific metabolites into and out of the organelles. In bacteria, the electrochemical proton gradient drives more than ATP synthesis and transport processes; as a store of directly usable energy, it also drives the rapid rotation of the bacterial flagellum, which enables the bacterium to swim.

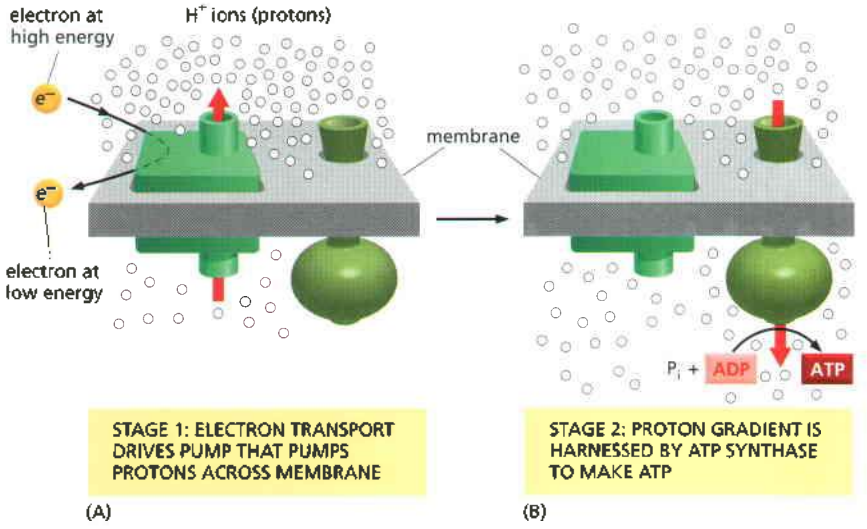


Figure 14–1 Harnessing energy for life. (A) The essential requirements for chemiosmosis are a membrane—in which are embedded a pump protein and an ATP synthase, plus a source of high-energy electrons (e^-). The protons (H^+) shown are freely available from water molecules. The pump harnesses the energy of electron transfer (details not shown here) to pump protons, creating an electrochemical proton gradient across the membrane. (B) This proton gradient serves as an energy store that can be used to drive ATP synthesis by the ATP synthase enzyme. The *red arrow* shows the direction of proton movement at each stage.

The entire set of proteins in the membrane, together with the small molecules involved in the orderly sequence of electron transfers, is called an **electron-transport chain**. The mechanism of electron transport is analogous to an electric cell driving a current through a set of electric pumps. However, in biological systems, electrons are carried between one site and another not by conducting wires, but by diffusible molecules that can pick up electrons at one location and deliver them to another. For mitochondria, the first of these electron carriers is NAD^+, which takes up two electrons (plus an H^+) to become NADH, a water-soluble small molecule that ferries electrons from the sites where food molecules are degraded to the inner mitochondrial membrane.

Figure 14–3 compares the electron-transport processes in mitochondria, which convert energy from chemical fuels, with those in chloroplasts, which convert energy from sunlight. In the mitochondrion, electrons—which have been released from a carbohydrate food molecule in the course of its degradation to CO_2—are transferred through the membrane by a chain of electron carriers, finally reducing oxygen gas (O_2) to form water. The free energy released as the electrons flow down this path from a high-energy state to a low-energy state drives a series of three H^+ pumps in the inner mitochondrial membrane, and it is the third H^+ pump in the series that catalyzes the transfer of the electrons to O_2 (see Figure 14–3A).

Although the chloroplast can be described in similar terms, and several of its main components are similar to those of the mitochondrion, the chloroplast membrane contains some crucial components not found in the mitochondrial membrane. Foremost among these are the *photosystems*, where the green pigment chlorophyll captures light energy and harnesses it to drive the transfer of electrons, much as photocells in solar panels absorb light energy and use it to drive an electric current. The electron-motive force generated by the chloroplast photosystems drives electron transfer in the direction opposite to that in mitochondria: electrons are taken from water to produce O_2, and they are donated to CO_2 (via NADPH, a compound closely related to NADH) to synthesize carbohydrate. Thus, the chloroplast generates O_2 and carbohydrate, whereas the mitochondrion consumes them (see Figure 14–3B).

It is thought that the energy-converting organelles of eucaryotes evolved from procaryotes that were engulfed by primitive eucaryotic cells and developed a symbiotic relationship with them (discussed in Chapter 12). This hypothesis explains why mitochondria and chloroplasts contain their own DNA, which codes for some of their proteins. Since their initial uptake by a host cell, these

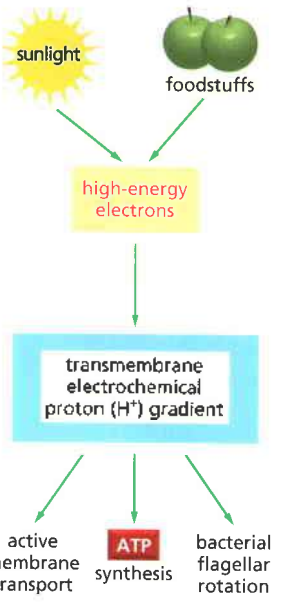


Figure 14–2 Chemiosmotic coupling. Energy from sunlight or the oxidation of foodstuffs is first used to create an electrochemical proton gradient across a membrane. This gradient serves as a versatile energy store that drives energy-requiring reactions in mitochondria, chloroplasts, and bacteria.

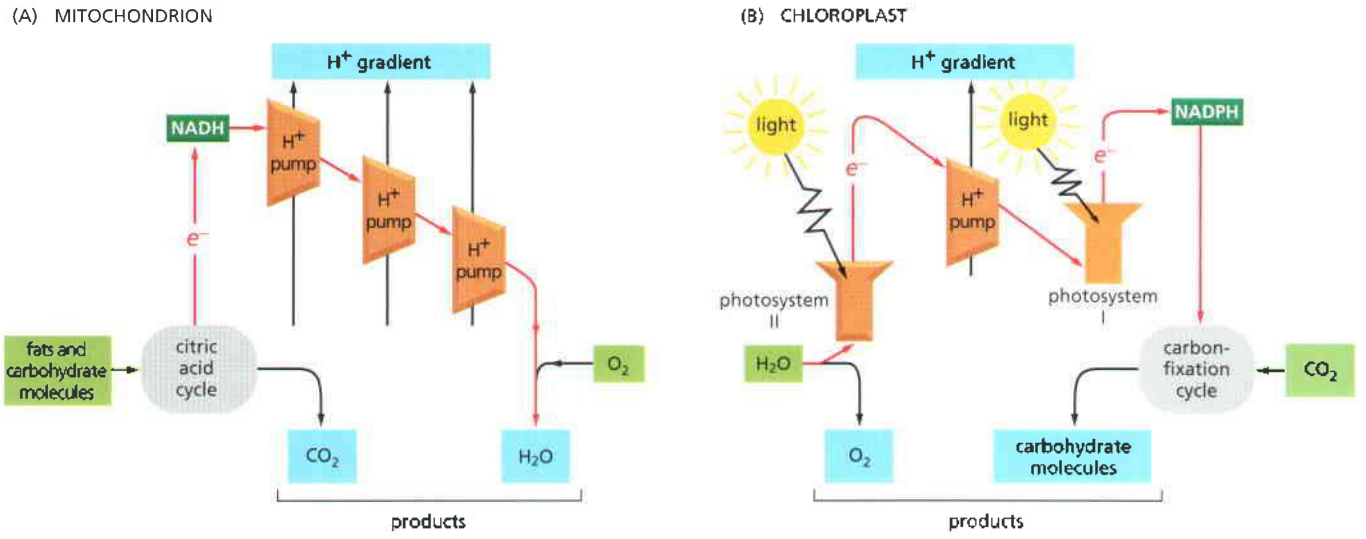


Figure 14–3 **Electron-transport processes.** (A) The mitochondrion converts energy from chemical fuels. (B) The chloroplast converts energy from sunlight. Inputs are *light green*, products are *blue*, and the path of electron flow is indicated by *red arrows*. Each of the protein complexes *(orange)* is embedded in a membrane. Note that the electron-motive force generated by the two chloroplast photosystems enables the chloroplast to drive electron transfer from H_2O to carbohydrate, and that this is *opposite* to the energetically favorable direction of electron transfer in a mitochondrion. Thus, whereas carbohydrate molecules and O_2 are inputs for the mitochondrion, they are products of the chloroplast.

organelles have lost much of their own genomes and become heavily dependent on proteins that are encoded by genes in the nucleus, synthesized in the cytosol, and then imported into the organelle. Conversely, the host cells have become dependent on these organelles for much of the ATP they need for biosyntheses, ion pumping, and movement; they have also become dependent on selected biosynthetic reactions that occur inside these organelles.

THE MITOCHONDRION

Mitochondria occupy a substantial portion of the cytoplasmic volume of eucaryotic cells, and they have been essential for the evolution of complex animals. Without mitochondria, present-day animal cells would have to depend on anaerobic glycolysis for all of their ATP. When glycolysis converts glucose to pyruvate, it releases only a small fraction of the total free energy that is potentially available from glucose oxidation. In mitochondria, the metabolism of sugars is completed: the pyruvate is imported into the mitochondrion and oxidized by O_2 to CO_2 and H_2O. This allows 15 times more ATP to be made than is produced by glycolysis alone.

Mitochondria are usually depicted as stiff, elongated cylinders with a diameter of 0.5–1 μm, resembling bacteria. Time-lapse microcinematography of living cells, however, shows that mitochondria are remarkably mobile and plastic organelles, constantly changing their shape (Figure 14–4) and even fusing with one another and then separating again. As they move about in the cytoplasm, they often seem to be associated with microtubules (Figure 14–5), which can determine the unique orientation and distribution of mitochondria in different types of cells. Thus, the mitochondria in some cells form long moving filaments or chains. In others they remain fixed in one position where they provide ATP directly to a site of unusually high ATP consumption—packed between adjacent myofibrils in a cardiac muscle cell, for example, or wrapped tightly around the flagellum in a sperm (Figure 14–6).

Mitochondria are large enough to be seen in the light microscope, and they were first identified during the nineteenth century. Real progress in understanding their function, however, depended on procedures developed in 1948 for isolating intact mitochondria. For technical reasons, many of these biochemical

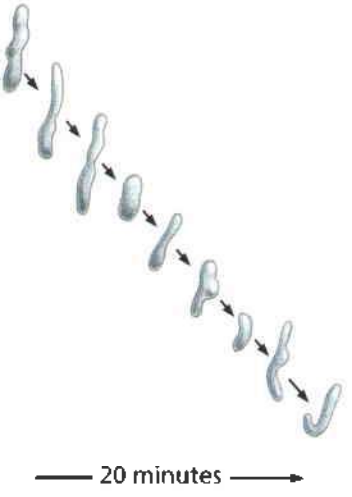


Figure 14–4 **Mitochondrial plasticity.** Rapid changes of shape are often observed when an individual mitochondrion is followed in a living cell.

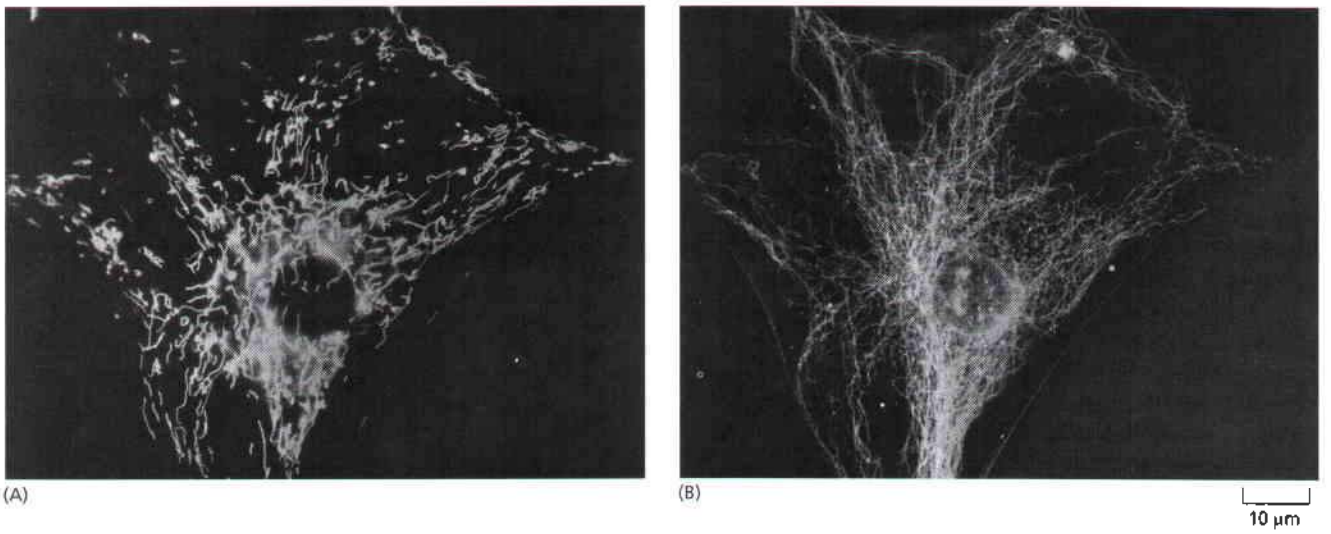


Figure 14–5 The relationship between mitochondria and microtubules. (A) A light micrograph of chains of elongated mitochondria in a living mammalian cell in culture. The cell was stained with a fluorescent dye (rhodamine 123) that specifically labels mitochondria in living cells. (B) An immunofluorescence micrograph of the same cell stained (after fixation) with fluorescent antibodies that bind to microtubules. Note that the mitochondria tend to be aligned along microtubules. (Courtesy of Lan Bo Chen.)

studies were performed with mitochondria purified from liver; each liver cell contains 1000–2000 mitochondria, which in total occupy about one-fifth of the cell volume.

The Mitochondrion Contains an Outer Membrane, an Inner Membrane, and Two Internal Compartments

Each mitochondrion is enclosed by two highly specialized membranes, which have very different functions. Together they create two separate mitochondrial compartments: the internal **matrix** and a much narrower **intermembrane space**. If purified mitochondria are gently disrupted and then fractionated into separate components (**Figure 14–7**), the biochemical composition of each of the two membranes and of the spaces enclosed by them can be determined. Each contains a unique collection of proteins. Most of these 1000 or so different mitochondrial proteins are encoded in the nucleus and imported into the mitochondrion from the cytoplasm by specialized protein translocases of the outer (TOM, translocase of the outer membrane) and inner (TIM) mitochondrial membrane (discussed in Chapter 12).

The **outer membrane** contains many *porin* molecules, a type of transport protein that forms large aqueous channels through the lipid bilayer (discussed in Chapter 11). This membrane thus resembles a sieve that is permeable to all molecules of 5000 daltons or less, including small proteins. Such molecules can enter the intermembrane space, but most of them cannot pass the impermeable inner membrane. Thus, whereas the intermembrane space is chemically equivalent to the cytosol with respect to the small molecules it contains, the matrix contains a highly selected set of these molecules.

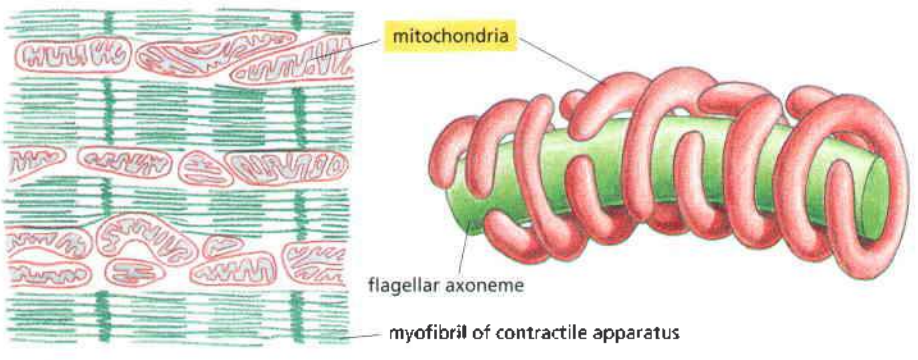


Figure 14–6 Localization of mitochondria near sites of high ATP utilization in cardiac muscle and a sperm tail. Cardiac muscle (A) in the wall of the heart is the most heavily used muscle in the body, and its continual contractions require a reliable energy supply. It has limited built in energy stores and has to depend on a steady supply of ATP from the copious mitochondria aligned close to the contractile myofibrils (see p. 1031). During the development of the flagellum of the sperm tail (B), microtubules wind helically around the axoneme, where they are thought to help localize the mitochondria in the tail; these microtubules then disappear, and the mitochondria fuse with one another to create the structure shown.

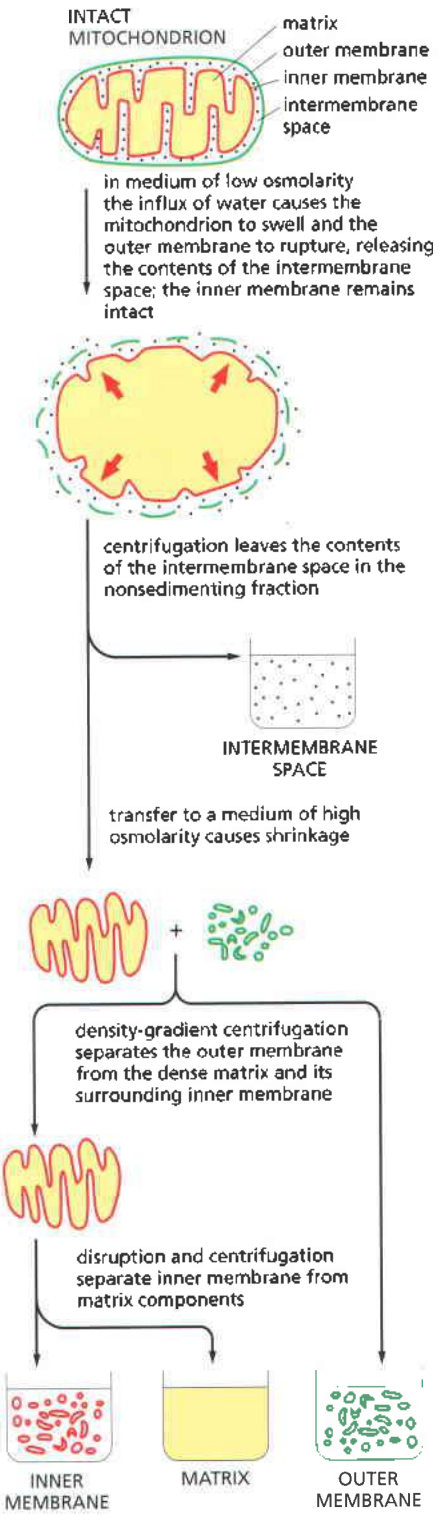


Figure 14–7 Biochemical fractionation of purified mitochondria into separate components. These techniques have made it possible to study the different proteins in each mitochondrial compartment. The method shown allows the processing of large numbers of mitochondria at the same time. It takes advantage of the fact that, in a solution of low osmotic strength, water flows into mitochondria and greatly expands the matrix space *(yellow)*. While the cristae of the inner membrane unfold to accommodate the expansion, the outer membrane—which has no folds—breaks, releasing a structure composed of only the inner membrane and the matrix.

As we explain in detail later, the major working part of the mitochondrion is the matrix and the **inner membrane** that surrounds it. The inner membrane is highly specialized. Its lipid bilayer contains a high proportion of the "double" phospholipid *cardiolipin*, which has four fatty acids rather than two and may help to make the membrane especially impermeable to ions (see Figure 14–65). This membrane also contains a variety of transport proteins that make it selectively permeable to those small molecules that are metabolized or required by the many mitochondrial enzymes concentrated in the matrix. The matrix enzymes include those that metabolize pyruvate and fatty acids to produce acetyl CoA and those that oxidize acetyl CoA in the *citric acid cycle*. The principal end products of this oxidation are CO_2, which is released from the cell as waste, and NADH, which is the main source of electrons for transport along the **respiratory chain**—the name given to the electron-transport chain in mitochondria. The enzymes of the respiratory chain are embedded in the inner mitochondrial membrane, and they are essential to the process of *oxidative phosphorylation*, which generates most of the animal cell's ATP.

As illustrated in **Figure 14–8**, the inner membrane is usually highly convoluted, forming a series of infoldings, known as **cristae**, that project into the matrix. These convolutions greatly increase the area of the inner membrane, so that in a liver cell, for example, it constitutes about one-third of the total cell membrane. The number of cristae is three times greater in the mitochondrion of a cardiac muscle cell than in the mitochondrion of a liver cell, presumably because of the greater demand for ATP in heart cells. There are also substantial differences in the mitochondrial enzymes of different cell types. In this chapter, we largely ignore these differences and focus instead on the enzymes and properties that are common to all mitochondria.

The Citric Acid Cycle Generates High-Energy Electrons

Mitochondria can use both pyruvate and fatty acids as fuel. Pyruvate comes from glucose and other sugars, whereas fatty acids come from fats. Both of these fuel molecules are transported across the inner mitochondrial membrane and are then converted to the crucial metabolic intermediate *acetyl CoA* by enzymes located in the mitochondrial matrix. The acetyl groups in acetyl CoA are then oxidized in the matrix via the **citric acid cycle**, described in Chapter 2. The cycle converts the carbon atoms in acetyl CoA to CO_2, which the cell releases as a waste product. Most importantly, this oxidation generates high-energy electrons, carried by the activated carrier molecules NADH and $FADH_2$ (**Figure 14–9**). These high-energy electrons are then transferred to the inner mitochondrial membrane, where they enter the electron-transport chain; the loss of electrons from NADH and $FADH_2$ also regenerates the NAD^+ and FAD that is needed for continued oxidative metabolism. **Figure 14–10** presents the entire sequence of reactions schematically.

A Chemiosmotic Process Converts Oxidation Energy into ATP

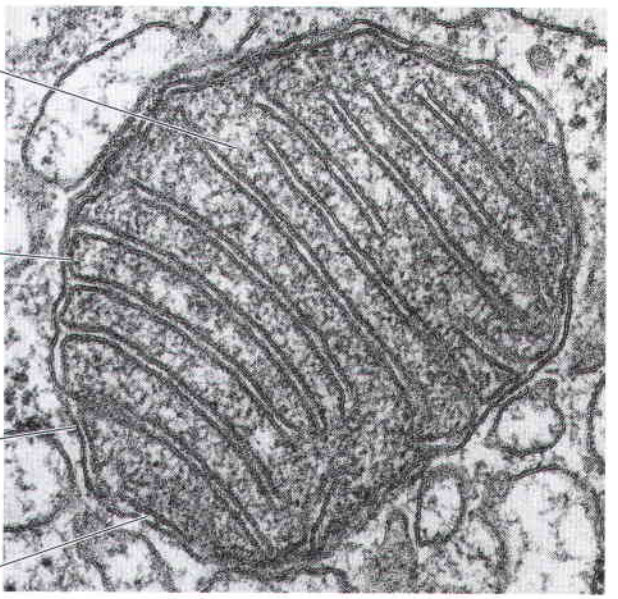
Although the citric acid cycle is considered to be part of aerobic metabolism, it does not itself use the oxygen. Only in the final catabolic reactions that take place on the inner mitochondrial membrane is molecular oxygen (O_2) directly

Matrix. This large internal space contains a highly concentrated mixture of hundreds of enzymes, including those required for the oxidation of pyruvate and fatty acids and for the citric acid cycle. The matrix also contains several identical copies of the mitochondrial DNA genome, special mitochondrial ribosomes, tRNAs, and various enzymes required for expression of the mitochondrial genes.

Inner membrane. The inner membrane is folded into numerous cristae, greatly increasing its total surface area. It contains proteins with three types of functions: (1) those that carry out the oxidation reactions of the electron-transport chain, (2) the ATP synthase that makes ATP in the matrix, and (3) transport proteins that allow the passage of metabolites into and out of the matrix. An electrochemical gradient of H^+, which drives the ATP synthase, is established across this membrane, so the membrane must be impermeable to ions and most small charged molecules.

Outer membrane. Because it contains a large channel-forming protein (a porin, VDAC), the outer membrane is permeable to all molecules of 5000 daltons or less. Other proteins in this membrane include enzymes involved in mitochondrial lipid synthesis and enzymes that convert lipid substrates into forms that are subsequently metabolized in the matrix, import receptors for mitochondrial proteins, and enzymatic machinery for division and fusion of the organelle.

Intermembrane space. This space contains several enzymes that use the ATP passing out of the matrix to phosphorylate other nucleotides.



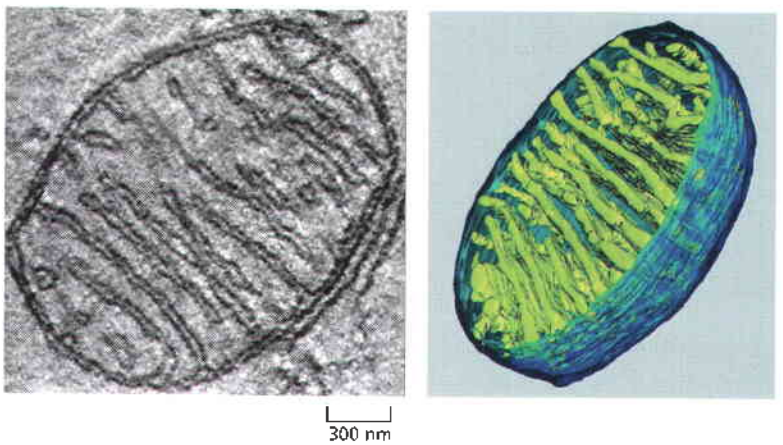


Figure 14–8 **The structure of a mitochondrion.** <CGAT> In the liver, an estimated 67% of the total mitochondrial protein is located in the matrix, 21% is located in the inner membrane, 6% in the outer membrane, and 6% in the intermembrane space. As indicated below, each of these four regions contains a special set of proteins that mediate distinct functions. (Large micrograph courtesy of Daniel S. Friend; small micrograph and three-dimensional reconstruction from T.G. Frey, C.W. Renken and G.A. Perkins, *Biochim. Biophys. Acta* 1555:196–203, 2002. With permission from Elsevier.)

consumed. Nearly all the energy available from burning carbohydrates, fats, and other foodstuffs in the earlier stages of their oxidation is initially saved in the form of high-energy electrons removed from substrates by NAD^+ and FAD. These electrons, carried by NADH and $FADH_2$, then combine with O_2 by means of the

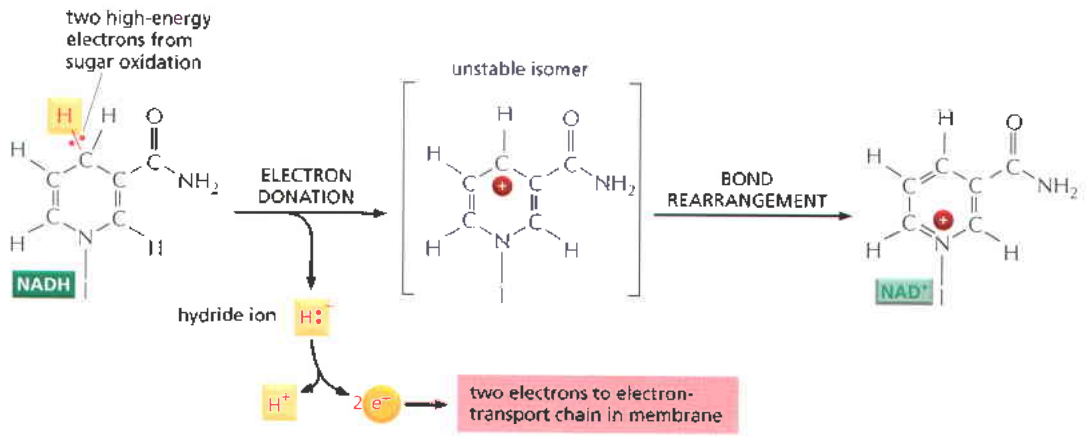


Figure 14–9 **How NADH donates electrons.** In this diagram, the high-energy electrons are shown as two *red dots* on a *yellow* hydrogen atom. A hydride ion (H^-, a hydrogen atom with an extra electron) is removed from NADH and is converted into a proton and two high-energy electrons: $H^- \rightarrow H^+ + 2e^-$. Only the ring that carries the electrons in a high-energy linkage is shown; for the complete structure and the conversion of NAD^+ back to NADH, see the structure of the closely related NADPH in Figure 2–60. Electrons are also carried in a similar way by $FADH_2$, whose structure is shown in Figure 2–83.

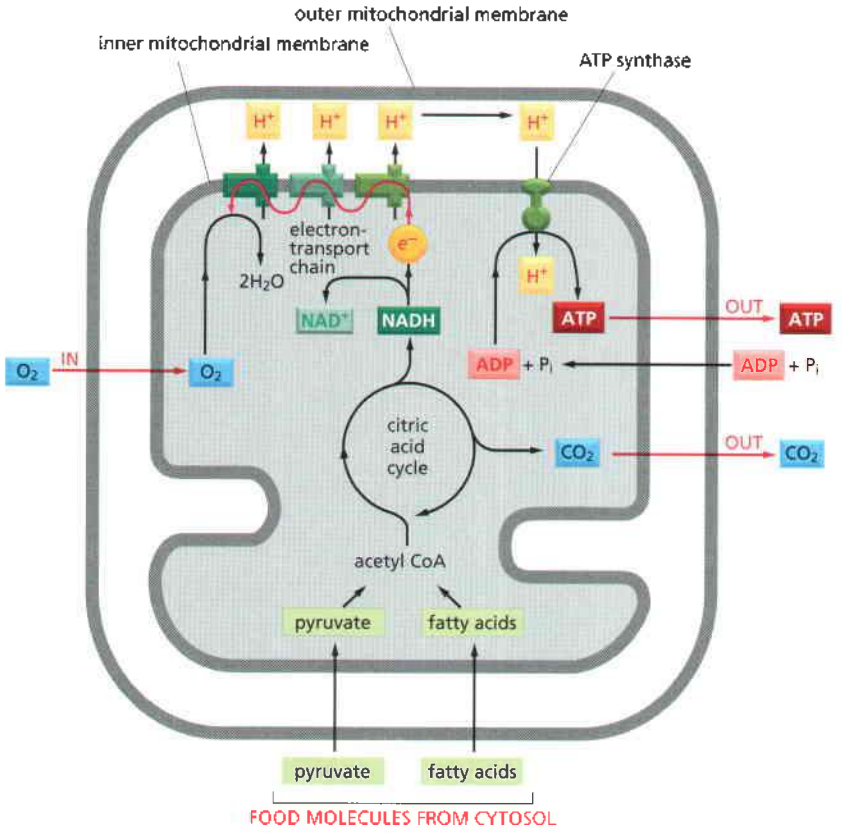


Figure 14–10 A summary of energy-generating metabolism in mitochondria. Pyruvate and fatty acids enter the mitochondrion *(bottom)* and are broken down to acetyl CoA. The acetyl CoA is then metabolized by the citric acid cycle, which reduces NAD^+ to NADH (and FAD to $FADH_2$, not shown). In the process of oxidative phosphorylation, high-energy electrons from NADH (and $FADH_2$) are then passed along the electron-transport chain in the inner membrane to oxygen (O_2). This electron transport generates a proton gradient across the inner membrane, which drives the production of ATP by ATP synthase (see Figure 14–1). The NADH generated by glycolysis in the cytosol also passes electrons to the respiratory chain (not shown). Since NADH cannot pass across the inner mitochondrial membrane, the electron transfer from cytosolic NADH must be accomplished indirectly by means of one of several "shuttle" systems that transport another reduced compound into the mitochondrion; after being oxidized, this compound is returned to the cytosol, where it is reduced by NADH again (see also Figure 14–32).

respiratory chain embedded in the inner mitochondrial membrane. The inner membrane harnesses the large amount of energy released to drive the conversion of ADP + P_i to ATP. For this reason, the term **oxidative phosphorylation** is used to describe this last series of reactions (**Figure 14–11**).

As previously mentioned, the generation of ATP by oxidative phosphorylation via the respiratory chain depends on a chemiosmotic process. When it was first proposed in 1961, this mechanism explained a long-standing puzzle in cell biology. Nonetheless, the idea was so novel that it was some years before enough supporting evidence accumulated to make it generally accepted.

In the remainder of this section we shall briefly describe the type of reactions that make oxidative phosphorylation possible, saving the details of the respiratory chain for later.

NADH Transfers its Electrons to Oxygen Through Three Large Respiratory Enzyme Complexes

Although the respiratory chain harvests energy by a different mechanism than that used in other catabolic reactions, the principle is the same. The energetically favorable reaction $H_2 + \frac{1}{2}\,O_2 \rightarrow H_2O$ is made to occur in many small steps, so that most of the energy released can be stored instead of being lost to the environment as heat. The hydrogen atoms are first separated into protons and electrons. The electrons pass through a series of electron carriers in the inner mitochondrial membrane. At several steps along the way, protons and electrons are transiently recombined. But only at the end of the electron-transport chain are the protons returned permanently, when they are used to neutralize the negative charges created by the final addition of the electrons to the oxygen molecule (**Figure 14–12**).

The process of electron transport begins when the hydride ion is removed from NADH (to regenerate NAD^+) and is converted into a proton and two electrons ($H^- \rightarrow H^+ + 2e^-$). The two electrons are passed to the first of the more than 15 different electron carriers in the respiratory chain. The electrons start with very high energy and gradually lose it as they pass along the chain. For the most

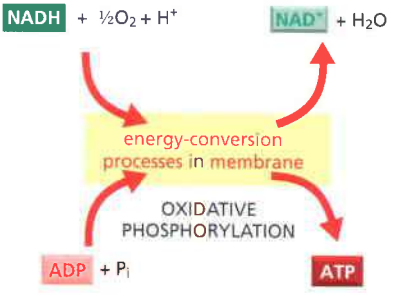


Figure 14–11 The major net energy conversion catalyzed by the mitochondrion. In this process of oxidative phosphorylation, the inner mitochondrial membrane serves as a device that changes one form of chemical bond energy to another, converting a major part of the energy of NADH (and $FADH_2$) oxidation into phosphate-bond energy in ATP.

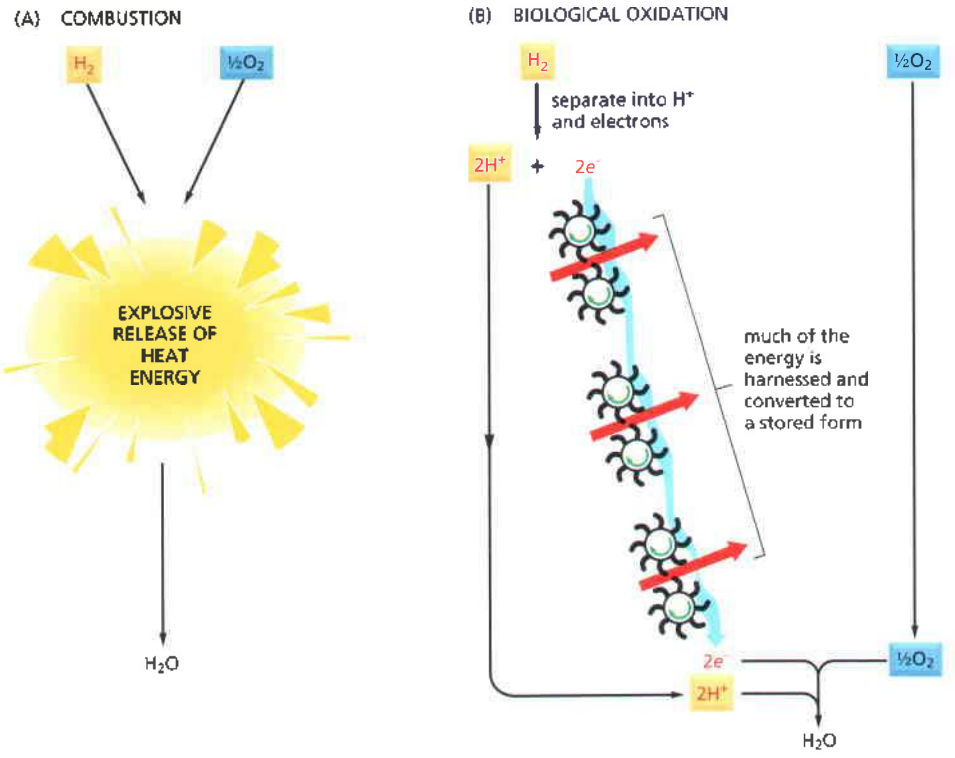


Figure 14–12 A comparison of biological oxidations with combustion. (A) Most of the energy would be released as heat if hydrogen were simply burned. (B) In biological oxidation by contrast, most of the released energy is stored in a form useful to the cell by means of the electron-transport chain in the inner mitochondrial membrane (the respiratory chain). The mitochondrion releases the rest of the oxidation energy as heat. In reality, the protons and electrons shown as being derived from H_2 are removed from hydrogen atoms that are covalently linked to NADH or $FADH_2$ molecules.

part, the electrons pass from one metal ion to another, each of these ions being tightly bound to a protein molecule that alters the electron affinity of the metal ion (discussed in detail later). Most of the proteins involved are grouped into three large *respiratory enzyme complexes*, each containing transmembrane proteins that hold the complex firmly in the inner mitochondrial membrane. Each complex in the chain has a greater affinity for electrons than its predecessor, and electrons pass sequentially from one complex to another until they are finally transferred to oxygen, which has the greatest affinity of all for electrons.

As Electrons Move Along the Respiratory Chain, Energy Is Stored as an Electrochemical Proton Gradient Across the Inner Membrane

The close association of the electron carriers with protein molecules makes oxidative phosphorylation possible. The proteins guide the electrons along the respiratory chain so that the electrons move sequentially from one enzyme complex to another. The transfer of electrons is coupled to oriented H^+ uptake and release, as well as to allosteric changes in energy-converting protein pumps. The net result is the pumping of H^+ across the inner membrane—from the matrix to the intermembrane space, driven by the energetically favorable flow of electrons. This movement of H^+ has two major consequences:

1. It generates a pH gradient across the inner mitochondrial membrane, with the pH higher in the matrix than in the cytosol, where the pH is generally close to 7. (Since small molecules equilibrate freely across the outer membrane of the mitochondrion, the pH in the intermembrane space is the same as in the cytosol.)
2. It generates a voltage gradient *(membrane potential)* across the inner mitochondrial membrane, with the inside negative and the outside positive (as a result of the net outflow of positive ions).

The pH gradient (ΔpH) drives H^+ back into the matrix, thereby reinforcing the effect of the membrane potential (ΔV), which acts to attract any positive ion into the matrix and to push any negative ion out. Together, the ΔpH and the ΔV are said to constitute an **electrochemical proton gradient** (Figure 14–13). The

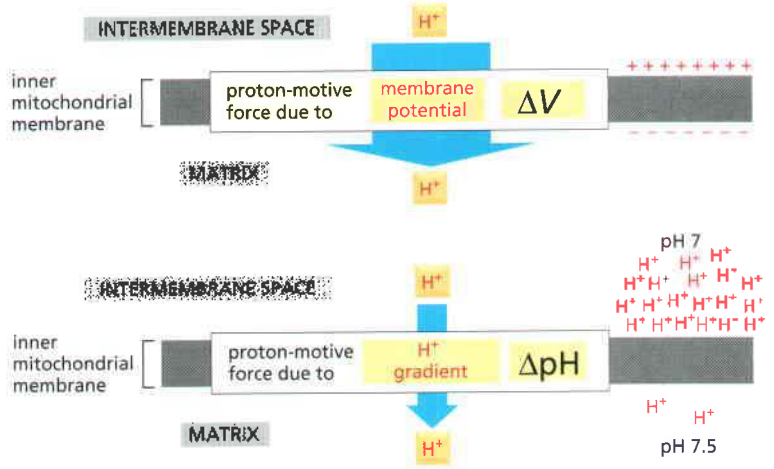


Figure 14–13 **The two components of the electrochemical proton gradient.** The total proton-motive force across the inner mitochondrial membrane consists of a large force due to the membrane potential (traditionally designated $\Delta\psi$ by experts, but designated ΔV in this text) and a smaller force due to the H^+ gradient (ΔpH). Both forces act to drive H^+ into the matrix.

electrochemical proton gradient exerts a **proton-motive force**, which can be measured in units of millivolts (mV). In a typical cell, the proton-motive force across the inner membrane of a respiring mitochondrion is about 180 to 190 mV (inside negative), and it is made up of a membrane potential of about 160 to 170 mV and a pH gradient of about 0.3 to 0.5 pH units (each ΔpH of 1 pH unit has an effect equivalent to a membrane potential of about 60 mV).

The Proton Gradient Drives ATP Synthesis

The electrochemical proton gradient across the inner mitochondrial membrane drives ATP synthesis in the critical process of oxidative phosphorylation (**Figure 14–14**). This is made possible by the membrane-bound enzyme **ATP synthase**, mentioned previously. This enzyme creates a hydrophilic pathway across the inner mitochondrial membrane that allows protons to flow down their electrochemical gradient. As these ions thread their way through the ATP synthase, they are used to drive the energetically unfavorable reaction between ADP and P_i that makes ATP (see Figure 2–27). The ATP synthase is of ancient origin; the same enzyme occurs in the mitochondria of animal cells, the chloroplasts of plants and algae, and in the plasma membrane of bacteria and archaea.

Figure 14–15 shows the structure of ATP synthase. Also called the F_0F_1 ATPase, it is a multisubunit protein with a mass of more than 500,000 daltons that works by rotary catalysis. A large enzymatic portion, shaped like a lollipop head and composed of a ring of 6 subunits, projects on the matrix side of the inner mitochondrial membrane. An elongated arm holds this head in place by tying it to a group of transmembrane proteins that form a "stator" in the membrane. This stator is in contact with a "rotor" composed of a ring of 10 to 14 identical transmembrane protein subunits. As protons pass through a narrow channel formed at the stator–rotor contact, their movement causes the rotor ring to spin. This spinning also turns a stalk attached to the rotor (*blue* in Figure 14–15B), which is thereby made to turn rapidly inside the lollipop head. As a result, the energy of proton flow down a gradient has been converted into the

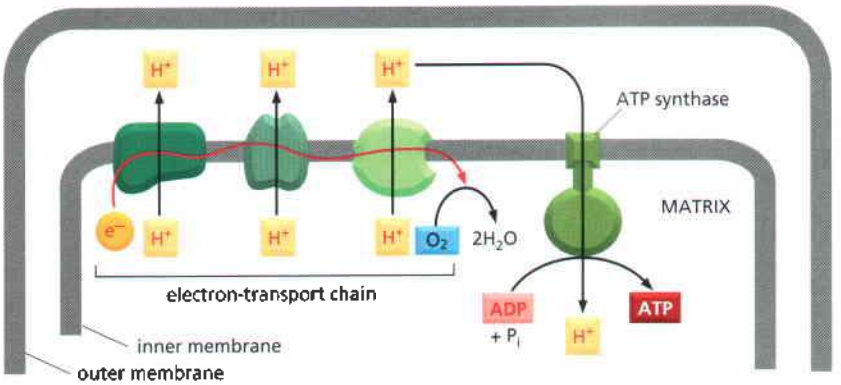


Figure 14–14 **The general mechanism of oxidative phosphorylation.** As a high-energy electron is passed along the electron-transport chain, some of the energy released drives the three respiratory enzyme complexes that pump H^+ out of the matrix. The resulting electrochemical proton gradient across the inner membrane drives H^+ back through the ATP synthase, a transmembrane protein complex that uses the energy of the H^+ flow to synthesize ATP from ADP and P_i in the matrix.

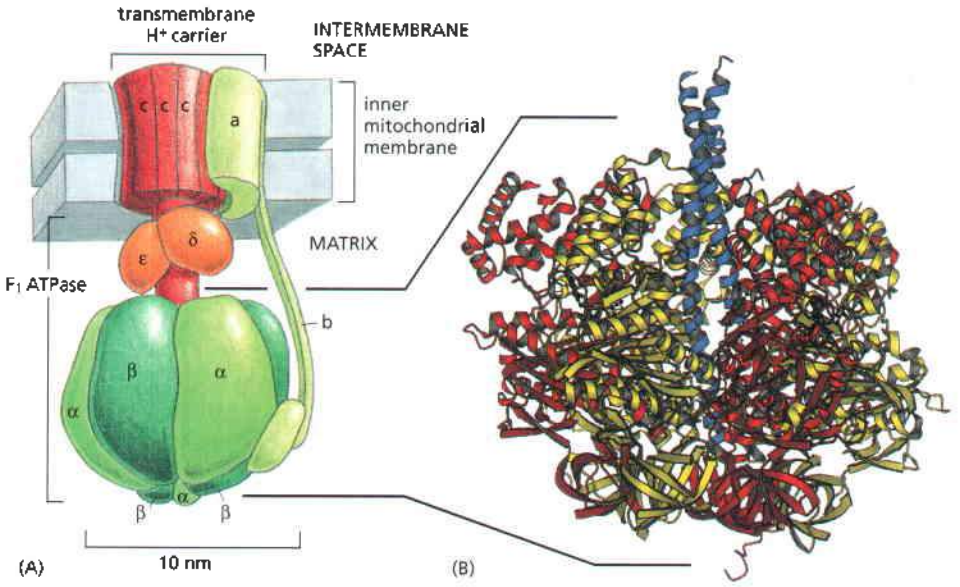


Figure 14–15 ATP synthase. <ATCG> <GAGA> (A) The enzyme is composed of a head portion, called the F_1 ATPase, and a transmembrane H^+ carrier, called F_0. Both F_1 and F_0 are formed from multiple subunits, as indicated. A rotating stalk is fixed to a rotor *(red)* formed by a ring of 10–14 c subunits in the membrane. The stator *(green)* is formed from transmembrane a subunits, tied to other subunits that create an elongated arm. This arm fixes the stator to a ring of 3α and 3β subunits that forms the head, which likewise cannot rotate. (B) The three-dimensional structure of the F_1 ATPase, determined by x-ray crystallography. This part of the ATP synthase derives its name from its ability to carry out the reverse of the ATP synthesis reaction—namely, the hydrolysis of ATP to ADP and P_i, when detached from the transmembrane portion. (B, courtesy of John Walker, from J.P. Abrahams et al., *Nature* 370:621–628, 1994. With permission from Macmillan Publishers Ltd.)

mechanical energy of two sets of proteins rubbing against each other: rotating stalk proteins pushing against a stationary ring of head proteins.

Three of the six subunits in the head contain binding sites for ADP and inorganic phosphate. These are driven to form ATP as mechanical energy is converted into chemical bond energy through the repeated changes in protein conformation that the rotating stalk creates. In this way, the ATP synthase is able to produce more than 100 molecules of ATP per second, generating 3 molecules of ATP per revolution. The number of proton-translocating subunits in the rotor is different in different ATP synthases, and it is this number that determines the number of protons that need to pass through this marvelous device to make each molecule of ATP (its "gear ratio", which is generally a non-integral number between 3 and 5).

The Proton Gradient Drives Coupled Transport Across the Inner Membrane

The electrochemical proton gradient drives other processes besides ATP synthesis. In mitochondria, many charged small molecules, such as pyruvate, ADP, and P_i, are pumped into the matrix from the cytosol, while others, such as ATP, must be moved in the opposite direction. Transporters that bind these molecules can couple their transport to the energetically favorable flow of H^+ into the mitochondrial matrix. Thus, for example, pyruvate and inorganic phosphate (P_i) are co-transported inward with H^+ as the H^+ moves into the matrix.

ADP and ATP are co-transported in opposite directions by a single transporter protein. Since an ATP molecule has one more negative charge than ADP, each nucleotide exchange results in a total of one negative charge being moved out of the mitochondrion. Thus, the voltage difference across the membrane drives this ADP–ATP co-transporter (**Figure 14–16**).

We have seen how, in eucaryotic cells, the electrochemical proton gradient across the inner mitochondrial membrane is used to drive both the formation of ATP and the transport of metabolites across the membrane. In bacteria, a similar gradient across the bacterial plasma membrane is harnessed to drive these two types of processes. In motile bacteria, this gradient also drives the rapid rotation of the bacterial flagellum, which propels the bacterium along (**Figure 14–17**).

Proton Gradients Produce Most of the Cell's ATP

As stated previously, glycolysis alone produces a net yield of 2 molecules of ATP for every molecule of glucose that is metabolized, and this is the total energy

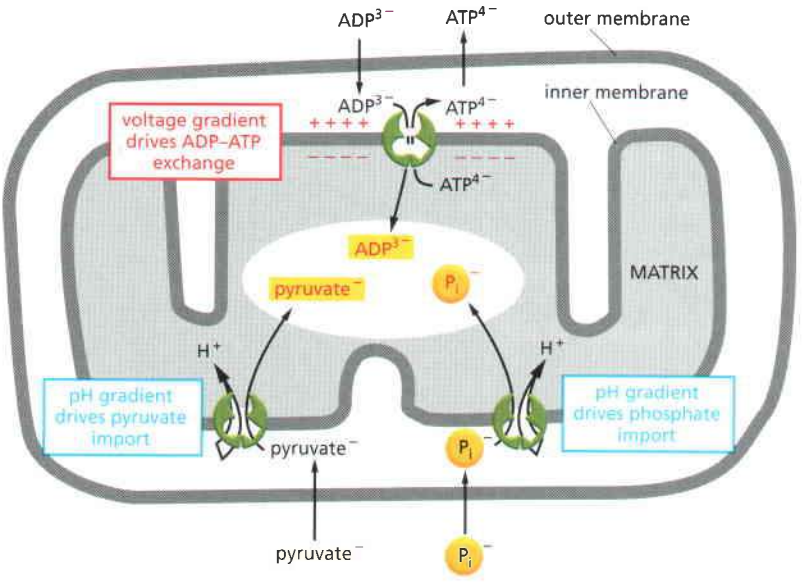


Figure 14–16 Some of the active transport processes driven by the electrochemical proton gradient across the inner mitochondrial membrane. Pyruvate, inorganic phosphate (P_i), and ADP are moved into the matrix, while ATP is pumped out. The charge on each of the transported molecules is indicated for comparison with the membrane potential, which is negative inside, as shown. The outer membrane is freely permeable to all of these compounds. The active transport of molecules across membranes by transporter proteins is discussed in Chapter 11.

yield for the fermentation processes that occur in the absence of O_2 (discussed in Chapter 2). During oxidative phosphorylation, each pair of electrons donated by the NADH produced in mitochondria is thought to provide energy for the formation of about 2.5 molecules of ATP, after subtracting the energy needed for transporting this ATP to the cytosol. Oxidative phosphorylation also produces 1.5 ATP molecules per electron pair from $FADH_2$, or from the NADH molecules produced by glycolysis in the cytosol. From the product yields of glycolysis and the citric acid cycle summarized in Table 14–1A, we can calculate that the complete oxidation of one molecule of glucose—starting with glycolysis and ending with oxidative phosphorylation—gives a net yield of about 30 ATPs.

In conclusion, the vast majority of the ATP produced from the oxidation of glucose in an animal cell is produced by chemiosmotic mechanisms in the mitochondrial membrane. Oxidative phosphorylation in the mitochondrion also produces a large amount of ATP from the NADH and the $FADH_2$ that is derived from the oxidation of fats (Table 14–1B; see also Figure 2–81).

Mitochondria Maintain a High ATP:ADP Ratio in Cells

Because of the carrier protein in the inner mitochondrial membrane that exchanges ATP for ADP, the ADP molecules produced by ATP hydrolysis in the

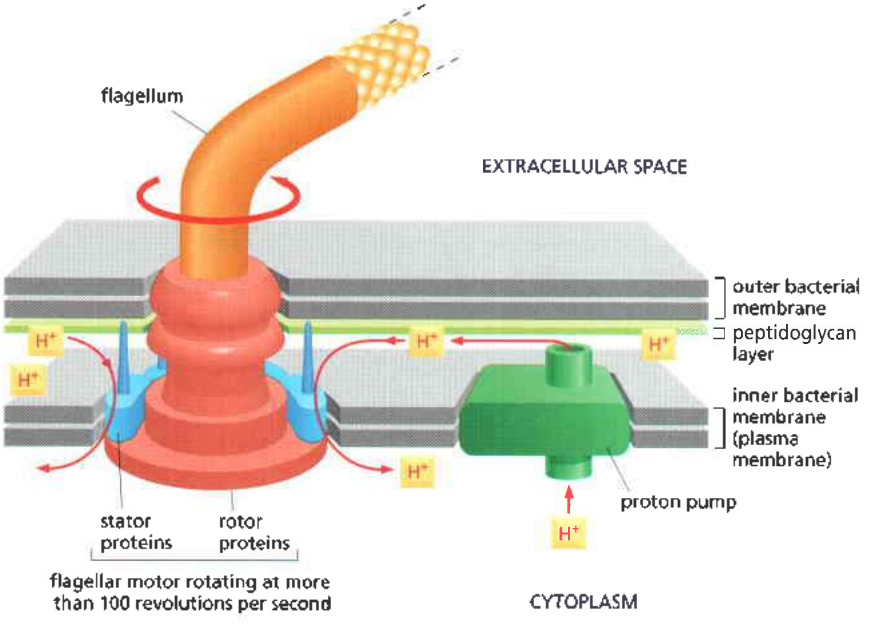


Figure 14–17 The rotation of the bacterial flagellum driven by H^+ flow. The flagellum is attached to a series of protein rings *(orange)*, which are embedded in the outer and inner membranes and rotate with the flagellum. The rotation is driven by a flow of protons through an outer ring of proteins (the stator) by mechanisms that may resemble those used by the ATP synthase.

Table 14–1 Product Yields from the Oxidation of Sugars and Fats

A. NET PRODUCTS FROM OXIDATION OF ONE MOLECULE OF GLUCOSE
In cytosol (glycolysis) 1 glucose → 2 pyruvate + 2 NADH + 2 ATP In mitochondrion (pyruvate dehydrogenase and citric acid cycle) 2 pyruvate → 2 acetyl CoA + 2 NADH 2 acetyl CoA → 6 NADH + 2 $FADH_2$ + 2 GTP
Net result in mitochondrion 2 pyruvate → 8 NADH + 2 $FADH_2$ + 2 GTP
B. NET PRODUCTS FROM OXIDATION OF ONE MOLECULE OF PALMITOYL COA (ACTIVATED FORM OF PALMITATE, A FATTY ACID)
In mitochondrion (fatty acid oxidation and citric acid cycle) 1 palmitoyl CoA → 8 acetyl CoA + 7 NADH + 7 $FADH_2$ 8 acetyl CoA → 24 NADH + 8 $FADH_2$ + 8 GTP
Net result in mitochondrion 1 palmitoyl CoA → 31 NADH + 15 $FADH_2$ + 8 GTP

cytosol rapidly enter mitochondria for recharging, while the ATP molecules formed in the mitochondrial matrix by oxidative phosphorylation are rapidly pumped into the cytosol where they are needed. A typical ATP molecule in the human body shuttles out of a mitochondrion and back into it (as ADP) for recharging more than once per minute, and the cell maintains a concentration of ATP that is about 10 times higher than the concentration of ADP.

As discussed in Chapter 2, biosynthetic enzymes often drive energetically unfavorable reactions by coupling them to the energetically favorable hydrolysis of ATP (see Figure 2–59). The ATP pool is therefore used to drive cellular processes in much the same way that a battery can be used to drive electric engines. If the activity of the mitochondria is blocked, ATP levels fall and the cell's battery runs down; eventually, energetically unfavorable reactions are no longer driven, and the cell dies. The poison cyanide, which blocks electron transport in the inner mitochondrial membrane, causes death in exactly this way.

It might seem that cellular processes would stop only when the concentration of ATP reaches zero; but, in fact, life is more demanding: it depends on cells maintaining a concentration of ATP that is high compared with the concentrations of ADP and P_i. To explain why, we must consider some elementary principles of thermodynamics.

A Large Negative Value of ΔG for ATP Hydrolysis Makes ATP Useful to the Cell

In Chapter 2, we introduced the concept of free energy (*G*). The free-energy change for a reaction, ΔG, determines whether this reaction will occur in a cell. We showed on p. 76 that the ΔG for a given reaction can be written as the sum of two parts: the first, called the *standard free-energy change*, $\Delta G°$, depends on the intrinsic characters of the reacting molecules; the second depends on their concentrations. For the simple reaction A → B,

$$\Delta G = \Delta G° + RT \ln \frac{[B]}{[A]}$$

where [A] and [B] denote the concentrations of A and B, and ln is the natural logarithm. $\Delta G°$ is therefore only a reference value that equals the value of ΔG when the molar concentrations of A and B are equal ($\ln 1 = 0$).

In Chapter 2, ATP was described as the major "activated energy carrier molecule" in cells. The large, favorable free-energy change (large negative ΔG) for its hydrolysis is used, through *coupled reactions*, to drive other chemical reactions that would otherwise not occur (see pp. 79–87). The ATP hydrolysis

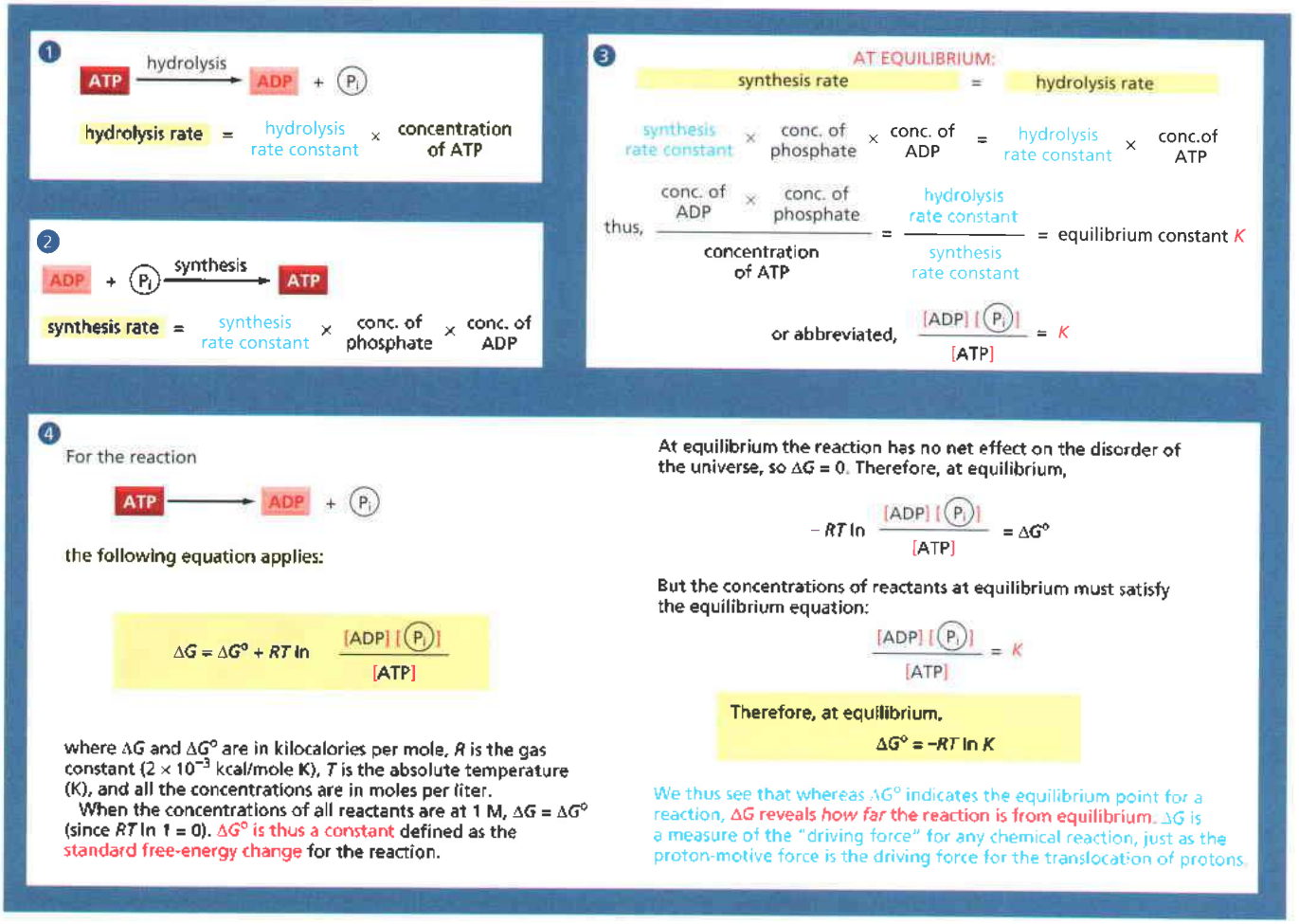


Figure 14–18 The basic relationship between free-energy changes and equilibrium in the ATP hydrolysis reaction. The rate constants in boxes 1 and 2 are determined from experiments in which product accumulation is measured as a function of time. The equilibrium constant shown here, *K*, is in units of moles per liter. (See Panel 2–7, pp. 118–119, for a discussion of free energy and see Figure 3–43 for a discussion of the equilibrium constant.)

reaction produces two products, ADP and inorganic phosphate (P_i); it is therefore of the type A → B + C, where, as described in **Figure 14–18**,

$$\Delta G = \Delta G^\circ + RT \ln \frac{[B][C]}{[A]}$$

When ATP is hydrolyzed to ADP and P_i under the conditions that normally exist in a cell, the free-energy change is roughly −11 to −13 kcal/mole (−46 to −54 kJ/mole). This extremely favorable ΔG depends on having a high concentration of ATP in the cell compared with the concentration of ADP and P_i. When ATP, ADP, and P_i are all present at the same concentration of 1 mole/liter (so-called standard conditions), the ΔG for ATP hydrolysis is the standard free-energy change (ΔG°), which is only −7.3 kcal/mole (−30.5 kJ/mol). At much lower concentrations of ATP relative to ADP and P_i, ΔG becomes zero. At this point, the rate at which ADP and P_i will join to form ATP will be equal to the rate at which ATP hydrolyzes to form ADP and P_i. In other words, when $\Delta G = 0$, the reaction is at *equilibrium* (see Figure 14–18).

It is ΔG, not ΔG°, that indicates how far a reaction is from equilibrium and determines whether it can be used to drive other reactions. Because the efficient conversion of ADP to ATP in mitochondria maintains such a high concentration of ATP relative to ADP and P_i, the ATP-hydrolysis reaction in cells is kept very far from equilibrium and ΔG is correspondingly very negative. Without this large disequilibrium, ATP hydrolysis could not be used to direct the reactions of the cell; for example, many biosynthetic reactions would run backward rather than forward at low ATP concentrations.

ATP Synthase Can Function in Reverse to Hydrolyze ATP and Pump H^+

In addition to harnessing the flow of H^+ down an electrochemical proton gradient to make ATP, the ATP synthase can work in reverse: it can use the energy of ATP hydrolysis to pump H^+ across the inner mitochondrial membrane (**Figure 14–19**). It thus acts as a *reversible coupling device*, interconverting electrochemical proton gradient and chemical bond energies. The direction of action at any instant depends on the balance between the steepness of the electrochemical proton gradient and the local ΔG for ATP hydrolysis, as we now explain.

The exact number of protons needed to make each ATP molecule depends on the number of subunits in the ring of transmembrane proteins that forms the base of the rotor (see Figure 14–15). However, to illustrate the principles involved, let us assume that one molecule of ATP is made by the ATP synthase for every 3 protons driven through it. Whether the ATP synthase works in its ATP-synthesizing or its ATP-hydrolyzing direction at any instant depends, in this case, on the exact balance between the favorable free-energy change for moving the three protons across the membrane into the matrix, ΔG_{3H^+} (which is less than zero), and the unfavorable free-energy change for ATP *synthesis* in the matrix, $\Delta G_{ATP\ synthesis}$ (which is greater than zero). As just discussed, the value of $\Delta G_{ATP\ synthesis}$ depends on the exact concentrations of the three reactants ATP, ADP, and P_i in the mitochondrial matrix (see Figure 14–18). The value of ΔG_{3H^+}, in contrast, is directly proportional to the value of the proton-motive force across the inner mitochondrial membrane. The following example will show how the balance between these two free-energy changes affects the ATP synthase.

As explained in the legend to Figure 14–19, a single H^+ moving into the matrix down an electrochemical gradient of 200 mV liberates 4.6 kcal/mole (19.2 kJ/mol) of free energy, while the movement of three protons liberates three times this much free energy (ΔG_{3H^+} = –13.8 kcal/mole; 57.7 kJ/mol). Thus, if the proton-motive force remains constant at 200 mV, the ATP synthase synthesizes ATP until a ratio of ATP to ADP and P_i is reached where $\Delta G_{ATP\ synthesis}$ is just equal to +13.8 kcal/mole (57.7 kJ/mol; here $\Delta G_{ATP\ synthesis} + \Delta G_{3H^+} = 0$). At this point there is no further net ATP synthesis or hydrolysis by the ATP synthase.

Suppose the energy-requiring reactions in the cytosol suddenly hydrolyze a large amount of ATP, causing the ATP:ADP ratio in the matrix to fall. Now the value of $\Delta G_{ATP\ synthesis}$ will decrease (see Figure 14–18), and ATP synthase will begin to synthesize ATP again to restore the original ATP:ADP ratio. Alternatively, if the proton-motive force drops suddenly and is then maintained at a constant 160 mV, ΔG_{3H^+} will change to –11.0 kcal/mole (–46 kJ/mol). As a result, ATP synthase will start hydrolyzing some of the ATP in the matrix until a new balance of ATP to ADP and P_i is reached (where $\Delta G_{ATP\ synthesis}$ = +11.0 kcal/mole, or +46 kJ/mol), and so on.

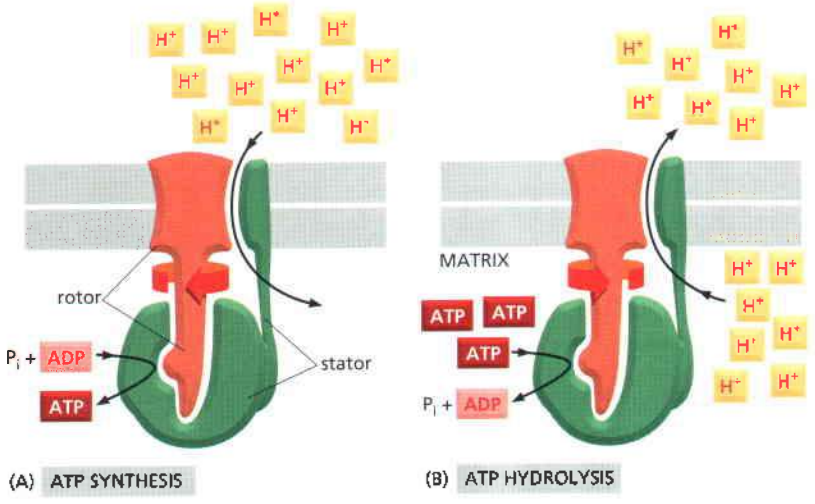


Figure 14–19 The ATP synthase is a reversible coupling device that can convert the energy of the electrochemical proton gradient into chemical-bond energy, or vice versa. <ATCG> The ATP synthase can either (A) synthesize ATP by harnessing the proton-motive force or (B) pump protons against their electrochemical gradient by hydrolyzing ATP. The direction of operation at any given instant depends on the net free-energy change (ΔG) for the coupled processes of H^+ translocation across the membrane and the synthesis of ATP from ADP and P_i. Measurement of the torque that the ATP synthase can produce when hydrolyzing ATP reveals that the synthase can pump 60 times more strongly than a diesel-engine of equal weight.

The free-energy change (ΔG) for ATP hydrolysis depends on the concentrations of the three reactants ATP, ADP, and P_i (see Figure 14–18); the ΔG for ATP synthesis is the negative of this value. The ΔG for proton translocation across the membrane is proportional to the proton-motive force. The conversion factor between them is the faraday. Thus, ΔG_{H^+} = –0.023 (proton-motive force), where ΔG_{H^+} is in kcal/mole and the proton-motive force is in mV. For an electrochemical proton gradient (proton-motive force) of 200 mV, ΔG_{H^+} = –4.6 kcal/mole (–19.2kJ/mole).

In many bacteria, the ATP synthase is routinely reversed in a transition between aerobic and anaerobic metabolism, as we shall see later. And the V-type ATPases that acidify organelles, which are architecturally similar to the ATP synthase, normally function in reverse (see Figure 13–36). Other membrane transport proteins that couple the transmembrane movement of an ion to ATP synthesis or hydrolysis share the same type of reversibility. Both the Na^+-K^+ pump and the Ca^{2+} pump described in Chapter 11, for example, normally hydrolyze ATP and use the energy released to move their specific ions across a membrane. If either of these pumps is exposed to an abnormally steep gradient of the ions it transports, however, it will act in reverse—synthesizing ATP from ADP and P_i instead of hydrolyzing it. Thus, the ATP synthase is by no means unique in its ability to convert the electrochemical energy stored in a transmembrane ion gradient directly into phosphate-bond energy in ATP.

Summary

The mitochondrion performs most cellular oxidations and produces the bulk of the animal cell's ATP. A mitochondrion is enclosed by two concentric membranes, and its major working part is the inner-most space (the matrix) and the inner membrane that surrounds it. The matrix contains a large variety of enzymes—including those that convert pyruvate and fatty acids to acetyl CoA and those that oxidize this acetyl CoA to CO_2 through the citric acid cycle. These oxidation reactions produce large amounts of NADH (and $FADH_2$). The electron-transport chain (respiratory chain) located in the inner mitochondrial membrane then harnesses the energy available from combining molecular oxygen with the reactive electrons carried by the NADH and $FADH_2$.

The respiratory chain uses energy derived from electron transport to pump H^+ out of the matrix to create a transmembrane electrochemical proton (H^+) gradient, which includes contributions from both a membrane potential and a pH difference. The large amount of free energy released when H^+ flows back into the matrix (across the inner membrane) provides the basis for ATP production in the matrix by a remarkable protein machine—the ATP synthase, a reversible coupling device between proton flows and ATP synthesis or hydrolysis. The transmembrane electrochemical gradient also drives the active transport of selected metabolites across the mitochondrial inner membrane, including an efficient ATP–ADP exchange between the mitochondrion and the cytosol that keeps the cell's ATP pool highly charged. The resulting high ratio of ATP to its hydrolysis products makes the free-energy change for ATP hydrolysis extremely favorable, allowing this hydrolysis reaction to drive a large number of energy-requiring processes throughout the cell.

ELECTRON-TRANSPORT CHAINS AND THEIR PROTON PUMPS

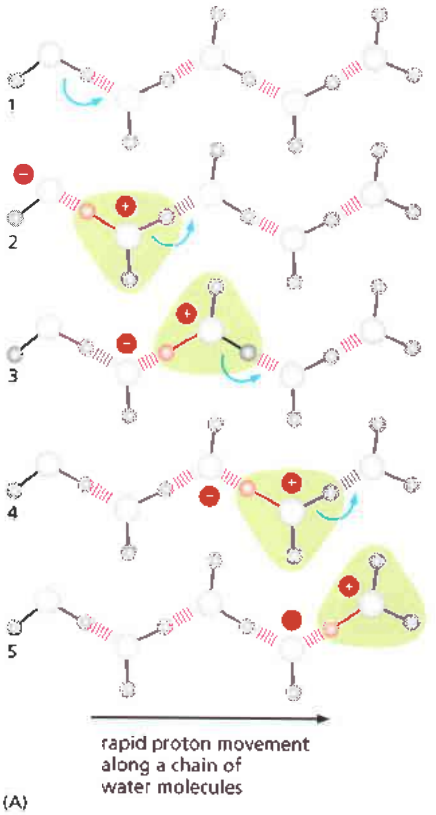
Having considered in general terms how a mitochondrion uses electron transport to create an electrochemical proton gradient, we now turn to the mechanisms that underlie this membrane-based energy-conversion process. In doing so, we also accomplish a larger purpose. As emphasized at the beginning of this chapter, mitochondria, chloroplasts, archaea, and bacteria use very similar chemiosmotic mechanisms. In fact, these mechanisms underlie the function of nearly all living organisms—including anaerobes that derive all their energy from electron transfers between two inorganic molecules.

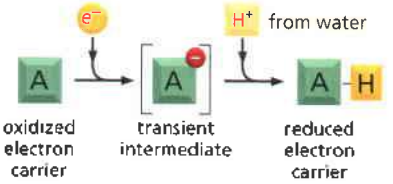
Our goal in this section is to explain how electron-transport process can pump protons across a membrane. We start with some of the basic principles on which this process depends.

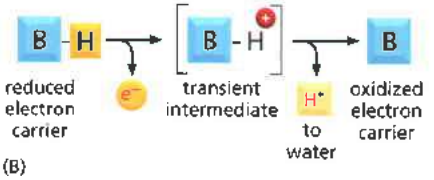
Protons Are Unusually Easy to Move

Although protons resemble other positive ions such as Na^+ and K^+ in normally requiring proteins to move them across membranes, in some respects they are

Figure 14–20 How protons behave in water. (A) Protons move very rapidly along a chain of hydrogen-bonded water molecules. In this diagram, proton jumps are indicated by *blue arrows*, and hydronium ions are indicated by *green shading*. As discussed in Chapter 2, naked protons rarely exist as such; they are instead associated with a water molecule in the form of a hydronium ion, H_3O^+. At a neutral pH (pH 7.0), the hydronium ions are present at a concentration of 10^{-7} M. However, for simplicity, we usually refer to this as an H^+ concentration of 10^{-7} M (see Panel 2–2, pp. 108–109). (B) Electron transfer can result in the transfer of entire hydrogen atoms, because protons are readily accepted from or donated to water inside cells. In this example, molecule A picks up an electron plus a proton when it is reduced, and B loses an electron plus a proton when it is oxidized.







unique. Hydrogen atoms are by far the most abundant type of atom in living organisms; they are plentiful not only in all carbon-containing biological molecules, but also in the water molecules that surround them. The protons in water are highly mobile, flickering through the hydrogen-bonded network of water molecules by rapidly dissociating from one water molecule to associate with its neighbor, as illustrated in Figure 14–20A. Protons are thought to move across a protein pump embedded in a lipid bilayer in a similar way: they transfer from one amino acid side chain to another, following a special channel through the protein.

Protons are also special with respect to electron transport. Whenever a molecule is reduced by acquiring an electron, the electron (e^-) brings with it a negative charge. In many cases, the addition of a proton (H^+) from water rapidly neutralizes this charge, so that the net effect of the reduction is to transfer an entire hydrogen atom, $H^+ + e^-$ (Figure 14–20B). Similarly, when a molecule is oxidized, a hydrogen atom removed from it can be readily dissociated into its constituent electron and proton—allowing the electron to be transferred separately to a molecule that accepts electrons, while the proton is passed to the water. Therefore, in a membrane in which electrons are being passed along an electron-transport chain, pumping protons from one side of the membrane to another can be relatively simple. The electron carrier merely needs to be arranged in the membrane in a way that causes it to pick up a proton from one side of the membrane when it accepts an electron, and to release the proton on the other side of the membrane as the electron is passed to the next carrier molecule in the chain (Figure 14–21).

The Redox Potential Is a Measure of Electron Affinities

In biochemical reactions, any electrons removed from one molecule are always passed to another, so that whenever one molecule is oxidized, another is reduced. Like any other chemical reaction, the tendency of such oxidation–reduction reactions, or **redox reactions**, to proceed spontaneously depends on the free-energy change (ΔG) for the electron transfer, which in turn depends on the relative affinities of the two molecules for electrons.

Because electron transfers provide most of the energy for living things, it is worth taking the time to understand them. As discussed in Chapter 2, acids and bases donate and accept protons (see Panel 2–2, pp. 108–109). Acids and bases exist in conjugate acid–base pairs, in which the acid is readily converted into the base by the loss of a proton. For example, acetic acid (CH_3COOH) is converted into its conjugate base (CH_3COO^-) in the reaction:

$$CH_3COOH \rightleftharpoons CH_3COO^- + H^+$$

In exactly the same way, pairs of compounds such as NADH and NAD^+ are called **redox pairs**, since NADH is converted to NAD^+ by the loss of electrons in the reaction:

$$NADH \rightleftharpoons NAD^+ + H^+ + 2e^-$$

NADH is a strong electron donor: because its electrons are held in a high-energy linkage, the free-energy change for passing its electrons to many other molecules is favorable (see Figure m14–9/14–9). It is difficult to form a high-energy linkage. Therefore its redox partner, NAD^+, is of necessity a weak electron acceptor.

We can measure the tendency to transfer electrons from any redox pair experimentally. All that is required is the formation of an electrical circuit linking a 1:1 (equimolar) mixture of the redox pair to a second redox pair that has been arbitrarily selected as a reference standard, so that we can measure the voltage difference between them (Panel 14–1, p. 830). This voltage difference is defined as the **redox potential**; as defined, electrons move spontaneously from a redox pair like NADH/NAD^+ with a low redox potential (a low affinity for electrons) to a redox pair like O_2/H_2O with a high redox potential (a high affinity for electrons). Thus, NADH is a good molecule for donating electrons to the respiratory chain, while O_2 is well suited to act as the "sink" for electrons at the end of the pathway. As explained in Panel 14–1, the difference in redox potential, $\Delta E_0'$, is a direct measure of the standard free-energy change (ΔG°) for the transfer of an electron from one molecule to another.

Electron Transfers Release Large Amounts of Energy

As just discussed, those pairs of compounds that have the most negative redox potentials have the weakest affinity for electrons and therefore contain carriers with the strongest tendency to donate electrons. Conversely, those pairs that have the most positive redox potentials have the strongest affinity for electrons and therefore contain carriers with the strongest tendency to accept electrons. A 1:1 mixture of NADH and NAD^+ has a redox potential of –320 mV, indicating that NADH has a strong tendency to donate electrons; a 1:1 mixture of H_2O and $\frac{1}{2}O_2$ has a redox potential of +820 mV, indicating that O_2 has a strong tendency to accept electrons. The difference in redox potential is 1.14 volts (1140 mV), which means that the transfer of each electron from NADH to O_2 under these standard conditions is enormously favorable, where ΔG° = –26.2 kcal/mole (110 kJ/mol), or twice this amount for the two electrons transferred per NADH molecule (see Panel 14–1). If we compare this free-energy change with that for the formation of the phosphoanhydride bonds in ATP, where ΔG° = –7.3 kcal/mole (–30.5 kJ/mol; see Figure 2–75), we see that the oxidization of one NADH molecule releases more than enough energy to synthesize several molecules of ATP from ADP and P_i.

Living systems could certainly have evolved enzymes that would allow NADH to donate electrons directly to O_2 to make water in the reaction:

$$2H^+ + 2e^- + \tfrac{1}{2}O_2 \rightarrow H_2O$$

But because of the huge free-energy drop, this reaction would proceed with almost explosive force and nearly all of the energy would be released as heat. Cells perform this reaction much more gradually by passing the high-energy electrons from NADH to O_2 via the many electron carriers in the electron-transport chain. Each successive carrier in the chain holds its electrons more tightly, so that the highly energetically favorable reaction $2H^+ + 2e^- + \frac{1}{2}O_2 \rightarrow H_2O$ occurs in many small steps. This stepwise process allows the cell to store nearly half of the energy that is released.

Spectroscopic Methods Identified Many Electron Carriers in the Respiratory Chain

Many of the electron carriers in the respiratory chain absorb visible light and change color when they are oxidized or reduced. In general, each has an absorption spectrum and reactivity that are distinct enough to allow its behavior to be traced spectroscopically, even in crude mixtures. It was therefore possible to purify these components long before their exact functions were known. Thus, the **cytochromes** were discovered in 1925 as compounds that undergo rapid oxidation and reduction in living organisms as disparate as bacteria, yeasts, and insects. By observing cells and tissues with a spectroscope, researchers identified three types of cytochromes by their distinctive absorption spectra and designated them as cytochromes *a*, *b*, and *c*. This nomenclature has survived, even though cells are now known to contain several cytochromes of each type and the classification into types is not functionally important.

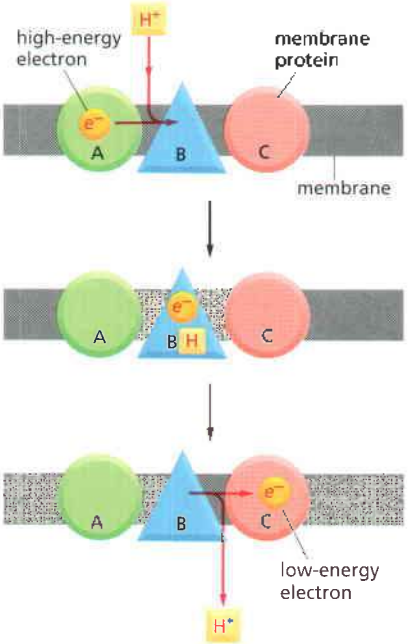
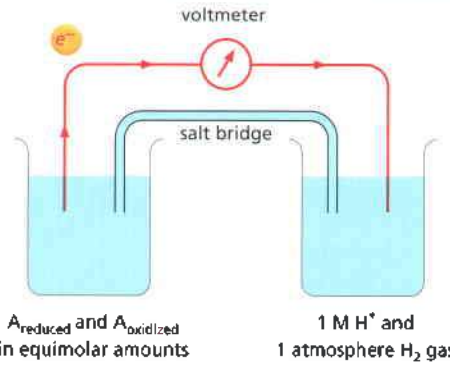


Figure 14–21 How protons can be pumped across membranes. As an electron passes along an electron-transport chain embedded in a lipid-bilayer membrane, it can bind and release a proton at each step. In this diagram, electron carrier B picks up a proton (H^+) from one side of the membrane when it accepts an electron (e^-) from carrier A; it releases the proton to the other side of the membrane when it donates its electron to carrier C.

HOW REDOX POTENTIALS ARE MEASURED



One beaker (*left*) contains substance A, with an equimolar mixture of the reduced ($A_{reduced}$) and oxidized ($A_{oxidized}$) members of its redox pair. The other beaker contains the hydrogen reference standard ($2H^+ + 2e^- \rightleftharpoons H_2$), whose redox potential is arbitrarily assigned as zero by international agreement. (A salt bridge formed from a concentrated KCl solution allows the ions K^+ and Cl^- to move between the two beakers, as required to neutralize the charges in each beaker when electrons flow between them.) The metal wire (*red*) provides a resistance-free path for electrons, and a voltmeter then measures the redox potential of substance A. If electrons flow from $A_{reduced}$ to H^+, as indicated here, the redox pair formed by substance A is said to have a negative redox potential. If they instead flow from H_2 to $A_{oxidized}$, the redox pair is said to have a positive redox potential.

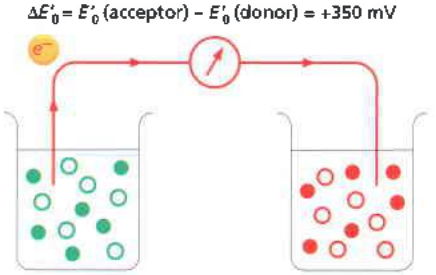
SOME STANDARD REDOX POTENTIALS AT pH7

By convention, the redox potential for a redox pair is designated E. For the standard state, with all reactants at a concentration of 1 M, including H^+, one can determine a standard redox potential, designated E_0. Since biological reactions occur at pH 7, biologists use a different standard state in which $A_{reduced} = A_{oxidized}$ and $H^+ = 10^{-7}$ M. This standard redox potential is designated E'_0. A few examples of special relevance to oxidative phosphorylation are given here.

redox reactions	redox potential E'_0
$NADH \rightleftharpoons NAD^+ + H^+ + 2e^-$	–320 mV
reduced ubiquinone $\rightleftharpoons$ oxidized ubiquinone $+ 2H^+ + 2e^-$	+30 mV
reduced cytochrome c $\rightleftharpoons$ oxidized cytochrome c $+ e^-$	+230 mV
$H_2O \rightleftharpoons \frac{1}{2}O_2 + 2H^+ + 2e^-$	+820 mV

CALCULATION OF ΔG° FROM REDOX POTENTIALS

$\Delta E'_0 = E'_0$ (acceptor) – E'_0 (donor) = +350 mV



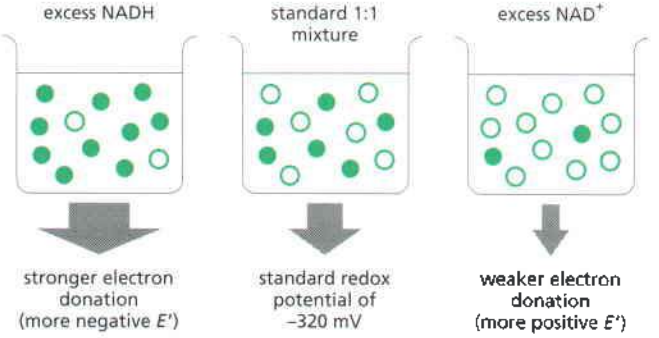
ΔG° = –8 kcal/mole

$\Delta G^\circ = -n(0.023)\,\Delta E'_0$, where n is the number of electrons transferred across a redox potential change of $\Delta E'_0$ millivolts (mV)

Example: The transfer of one electron from NADH to ubiquinone has a favorable ΔG° of –8.0 kcal/mole (–1.9 kJ/mole), whereas the transfer of one electron from ubiquinone to oxygen has an even more favorable ΔG° of –18.2 kcal/mole (–4.35 kJ/mole). The ΔG° value for the transfer of one electron from NADH to oxygen is the sum of these two values, –26.2 kcal/mole.

THE EFFECT OF CONCENTRATION CHANGES

The actual free-energy change for a reaction, ΔG, depends on the concentration of the reactants and generally is different from the standard free-energy change, ΔG°. The standard redox potentials are for a 1:1 mixture of the redox pair. For example, the standard redox potential of –320 mV is for a 1:1 mixture of NADH and NAD^+. But when there is an excess of NADH over NAD^+, electron transfer from NADH to an electron acceptor becomes more favorable. This is reflected by a more negative redox potential and a more negative ΔG for electron transfer.



The cytochromes constitute a family of colored proteins that are related by the presence of a bound *heme group*, whose iron atom changes from the ferric oxidation state (Fe^{3+}) to the ferrous oxidation state (Fe^{2+}) whenever it accepts an electron. The heme group consists of a *porphyrin ring* with a tightly bound iron atom held by four nitrogen atoms at the corners of a square (**Figure 14–22**). A similar porphyrin ring is responsible for the red color of blood and for the green color of leaves, being bound to iron in hemoglobin and to magnesium in chlorophyll, respectively.

In *iron–sulfur proteins*, a second major family of electron carriers, either two or four iron atoms are bound to an equal number of sulfur atoms and to cysteine side chains, forming an **iron–sulfur center** on the protein (**Figure 14–23**). There are more iron–sulfur centers than cytochromes in the respiratory chain. But their spectroscopic detection requires electron paramagnetic resonance (EPR) spectroscopy, and they are less completely characterized. Like the cytochromes, these centers carry one electron at a time.

The simplest of the electron carriers in the respiratory chain—and the only one that is not part of a protein—is a quinone (called *ubiquinone*, or *coenzyme Q*). A **quinone (Q)** is a small hydrophobic molecule that is freely mobile in the lipid bilayer and can pick up or donate either one or two electrons; upon reduction, it picks up a proton from the medium along with each electron it carries (**Figure 14–24**).

In addition to six different hemes linked to cytochromes, more than seven iron–sulfur centers, and ubiquinone, there are also two copper atoms and a flavin serving as electron carriers tightly bound to respiratory-chain proteins in the pathway from NADH to oxygen. This pathway involves more than 60 different proteins in all.

As we would expect, these electron carriers have higher and higher affinities for electrons (greater redox potentials) as the electrons move along the respiratory chain. The redox potentials have been fine-tuned during evolution by the binding of each electron carrier in a particular protein context, which can alter its normal affinity for electrons. However, because iron–sulfur centers have a relatively low affinity for electrons, they predominate in the early part of the respiratory chain; in contrast, the cytochromes predominate further down the chain, where a higher affinity for electrons is required.

The order of the individual electron carriers in the chain was determined by sophisticated spectroscopic measurements (**Figure 14–25**), and many of the proteins were initially isolated and characterized as individual polypeptides. A major advance in understanding the respiratory chain, however, was the later realization that most of the proteins are organized into three large enzyme complexes.

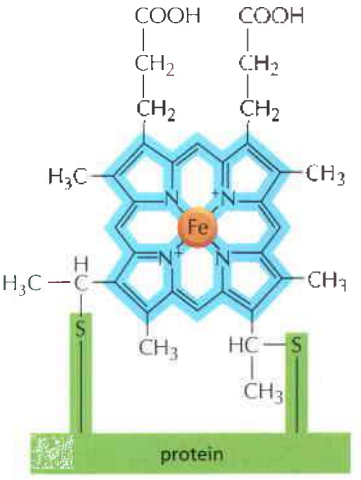


Figure 14–22 The structure of the heme group attached covalently to cytochrome *c*. The porphyrin ring is shown in *blue*. There are five different cytochromes in the respiratory chain. Because the hemes in different cytochromes have slightly different structures and are held by their respective proteins in different ways, each of the cytochromes has a different affinity for an electron.

The Respiratory Chain Includes Three Large Enzyme Complexes Embedded in the Inner Membrane

Membrane proteins are difficult to purify as intact complexes because they are insoluble in aqueous solutions, and some of the detergents required to solubilize them can destroy normal protein–protein interactions. In the early 1960s, however, researchers discovered that relatively mild ionic detergents, such as

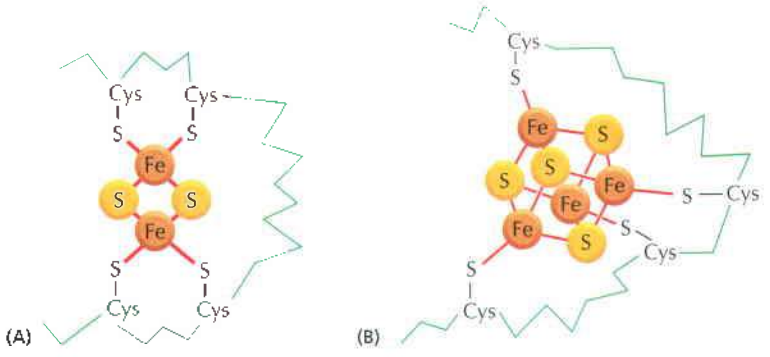


Figure 14–23 The structures of two types of iron–sulfur centers. (A) A center of the 2Fe2S type. (B) A center of the 4Fe4S type. Although they contain multiple iron atoms, each iron–sulfur center can carry only one electron at a time. There are more than seven different iron–sulfur centers in the respiratory chain.

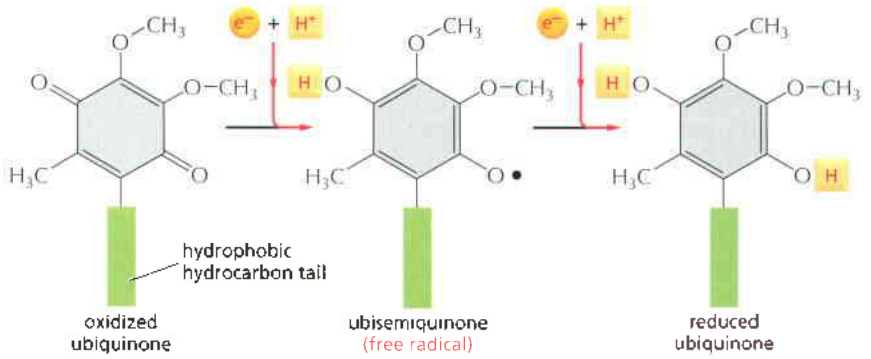


Figure 14–24 Quinone electron carriers. Ubiquinone in the respiratory chain picks up one H^+ from the aqueous environment for every electron it accepts, and it can carry either one or two electrons as part of a hydrogen atom *(yellow)*. When reduced ubiquinone donates its electrons to the next carrier in the chain, these protons are released. A long hydrophobic tail confines ubiquinone to the membrane and consists of 6–10 five-carbon isoprene units, the number depending on the organism. The corresponding electron carrier in the photosynthetic membranes of chloroplasts is plastoquinone, which is almost identical in structure. For simplicity, we refer to both ubiquinone and plastoquinone in this chapter as quinone (abbreviated as Q).

deoxycholate, can solubilize selected components of the inner mitochondrial membrane in their native form. This permitted the identification and purification of the three major membrane-bound **respiratory enzyme complexes** in the pathway from NADH to oxygen. Each of these purified complexes can be inserted into lipid bilayer vesicles and shown to pump protons across the bilayer as electrons pass through it. In the mitochondrion, the three complexes are asymmetrically oriented in the inner membrane, and they are linked in series as electron-transport-driven H^+ pumps that pump protons out of the matrix (Figure 14–26):

1. The **NADH dehydrogenase complex** (generally known as complex I) is the largest of the respiratory enzyme complexes, containing more than 40 polypeptide chains. It accepts electrons from NADH and passes them through a flavin and at least seven iron–sulfur centers to ubiquinone. Ubiquinone then transfers its electrons to a second respiratory enzyme complex, the cytochrome *b*-c_1 complex.
2. The **cytochrome *b*-c_1 complex** contains at least 11 different polypeptide chains and functions as a dimer. Each monomer contains three hemes bound to cytochromes and an iron–sulfur protein. The complex accepts electrons from ubiquinone and passes them on to cytochrome *c*, which carries its electron to the cytochrome oxidase complex.
3. **The cytochrome oxidase complex** also functions as a dimer; each monomer contains 13 different polypeptide chains, including two cytochromes and two copper atoms. The complex accepts one electron at a time from cytochrome *c* and passes them four at a time to oxygen.

The cytochromes, iron–sulfur centers, and copper atoms can carry only one electron at a time. Yet each NADH donates two electrons, and each O_2 molecule must receive four electrons to produce water. There are several electron-collecting and electron-dispersing points along the electron-transport chain that coordinate these changes in electron number. The most obvious of these is cytochrome oxidase.

An Iron–Copper Center in Cytochrome Oxidase Catalyzes Efficient O_2 Reduction

Because oxygen has a high affinity for electrons, it releases a large amount of free energy when it is reduced to form water. Thus, the evolution of cellular

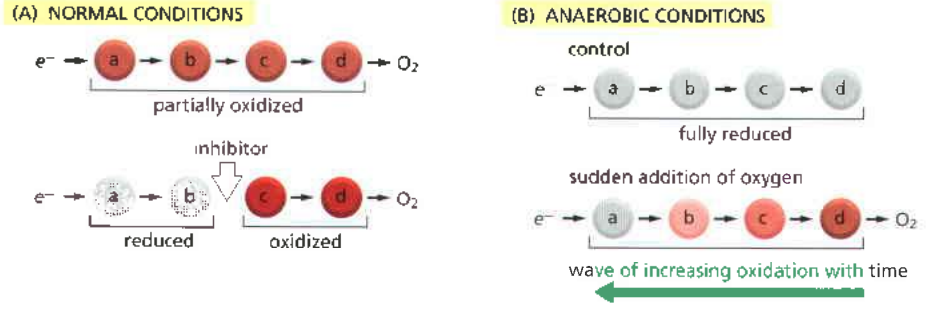


Figure 14–25 The general methods used to determine the path of electrons along an electron-transport chain. The extent of oxidation of electron carriers a, b, c, and d is continuously monitored by following their distinct spectra, which differ in their oxidized and reduced states. In this diagram, darker red indicates an increased degree of oxidation. (A) Under normal conditions, where oxygen is abundant, all carriers are in a partly oxidized state. The addition of a specific inhibitor causes the downstream carriers to become more oxidized *(red)* and the upstream carriers to become more reduced. (B) In the absence of oxygen, all carriers are in their fully reduced state *(gray)*. The sudden addition of oxygen converts each carrier to its partly oxidized form with a delay that is greatest for the most upstream carriers.

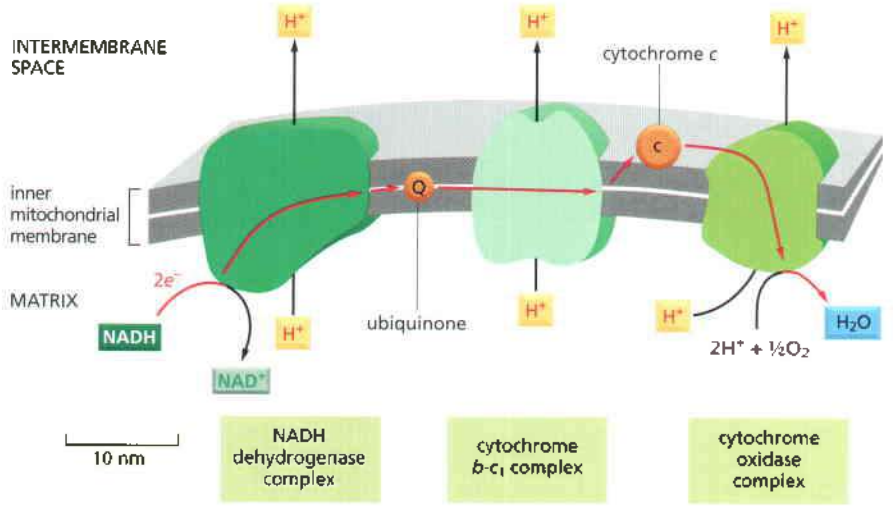


Figure 14–26 The path of electrons through the three respiratory enzyme complexes. The approximate size and shape of each complex are shown. During the transfer of electrons from NADH to oxygen *(red lines)*, ubiquinone and cytochrome *c* serve as mobile carriers that ferry electrons from one complex to the next. As indicated, protons are pumped across the membrane by each of the respiratory enzyme complexes.

respiration, in which O_2 is converted to water, enabled organisms to harness much more energy than can be derived from anaerobic metabolism. This is presumably why all higher organisms respire. The ability of biological systems to use O_2 in this way, however, requires a very sophisticated chemistry. We can tolerate O_2 in the air we breathe because it has trouble picking up its first electron; this fact has allowed cells to control its initial reaction through enzymatic catalysis. But once a molecule of O_2 has picked up one electron to form a superoxide radical (O_2^-), it becomes dangerously reactive and rapidly takes up an additional three electrons wherever it can find them. The cell can use O_2 for respiration only because cytochrome oxidase holds onto oxygen at a special bimetallic center, where it remains clamped between a heme-linked iron atom and a copper atom until it has picked up a total of four electrons. Only then can the two oxygen atoms of the oxygen molecule be safely released as two molecules of water (**Figure 14–27**).

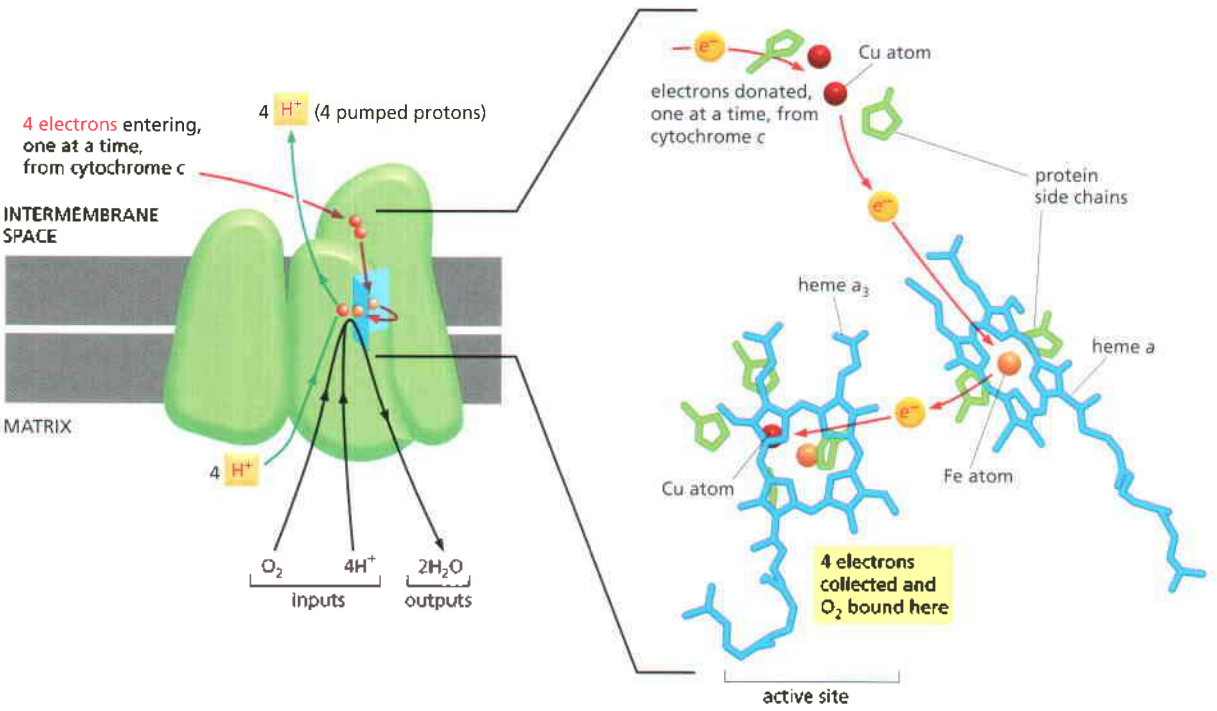


Figure 14–27 The reaction of O_2 with electrons in cytochrome oxidase. As indicated, the iron atom in heme *a* serves as an electron queuing point; this heme feeds four electrons into an O_2 molecule held at the bimetallic center active site, which is formed by the other heme-linked iron (in heme a_3) and a closely opposed copper atom. Note that four protons are pumped out of the matrix for each O_2 molecule that undergoes the reaction $4e^- + 4H^+ + O_2 \rightarrow 2H_2O$.

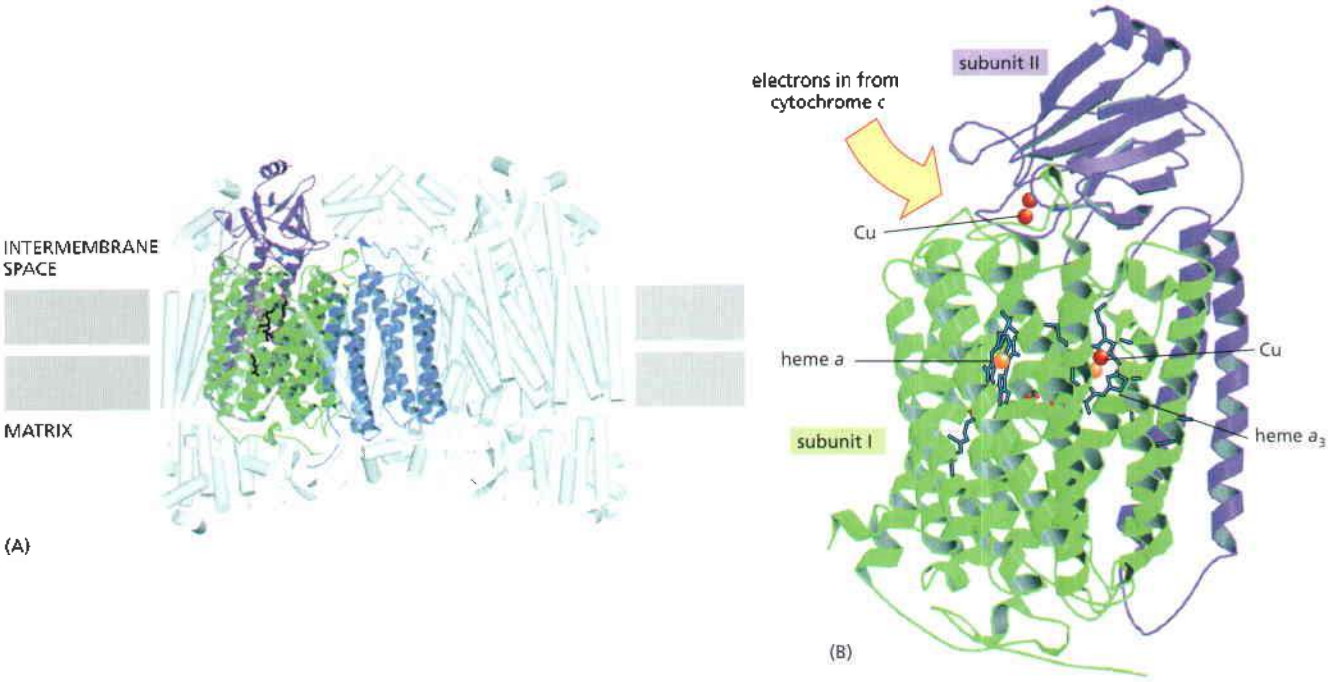


Figure 14–28 **The molecular structure of cytochrome oxidase.** This protein is a dimer formed from a monomer with 13 different protein subunits (monomer mass of 204,000 daltons). The three colored subunits are encoded by the mitochondrial genome, and they form the functional core of the enzyme. As electrons pass through this protein on the way to its bound O_2 molecule, they cause the protein to pump protons across the membrane (see Figure 14–27). (A) The entire protein is shown, positioned in the inner mitochondrial membrane. (B) The electron carriers are located in subunits I and II, as indicated.

The cytochrome oxidase reaction accounts for about 90% of the total oxygen uptake in most cells. This protein complex is therefore crucial for all aerobic life. Cyanide and azide are extremely toxic because they bind tightly to the cell's cytochrome oxidase complexes to stop electron transport, thereby greatly reducing ATP production.

Although the cytochrome oxidase in mammals contains 13 different protein subunits, most of these seem to have a subsidiary role, helping to regulate either the activity or the assembly of the three subunits that form the core of the enzyme. The complete structure of this large enzyme complex has been determined by x-ray crystallography, as illustrated in **Figure 14–28**. The atomic resolution structures, combined with mechanistic studies of the effect of precisely tailored mutations introduced into the enzyme by genetic engineering of the yeast and bacterial proteins, are revealing the detailed mechanisms of this finely tuned protein machine.

Electron Transfers in the Inner Mitochondrial Membrane Are Mediated by Electron Tunneling during Random Collisions

The two components that carry electrons between the three major enzyme complexes of the respiratory chain—ubiquinone and cytochrome *c*—diffuse rapidly in the plane of the inner mitochondrial membrane. The expected rate of random collisions between these mobile carriers and the more slowly diffusing enzyme complexes can account for the observed rates of electron transfer (each complex donates and receives an electron about once every 5–20 milliseconds).

The ordered transfer of electrons along the respiratory chain is due entirely to the specificity of the functional interactions between the components of the chain: each electron carrier is able to interact only with the carrier adjacent to it in the sequence shown in Figure 14–26, with no short circuits.

Electrons move between the molecules that carry them in biological systems not only by moving along covalent bonds within a molecule, but also by jumping across a gap as large as 2 nm. The jumps occur by electron "tunneling," a quantum-mechanical property that is critical for the processes we are discussing. Insulation prevents short circuits that would otherwise occur when an electron carrier with a low redox potential collides with a carrier with a high redox potential. This insulation seems to be provided by carrying an electron deep enough inside a protein to prevent its tunneling interactions with an inappropriate partner.

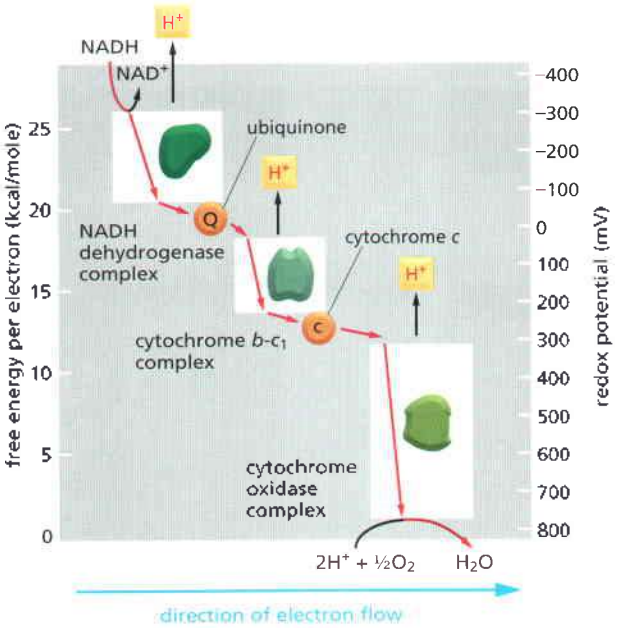


Figure 14–29 Redox potential changes along the mitochondrial electron-transport chain. The redox potential (designated E'_0) increases as electrons flow down the respiratory chain to oxygen. The standard free-energy change, ΔG°, for the transfer of each of the two electrons donated by an NADH molecule can be obtained from the left-hand ordinate ($\Delta G = -n(0.023)\,\Delta E'_0$, where n is the number of electrons transferred across a redox potential change of $\Delta E'_0$ mV). Electrons flow through a respiratory enzyme complex by passing in sequence through the multiple electron carriers in each complex. As indicated, part of the favorable free-energy change is harnessed by each enzyme complex to pump H^+ across the inner mitochondrial membrane. The NADH dehydrogenase and cytochrome b-c_1 complexes each pump two H^+ per electron, whereas the cytochrome oxidase complex pumps one.

Note that NADH is not the only source of electrons for the respiratory chain. The flavin $FADH_2$ is also generated by fatty acid oxidation (see Figure 2–81) and by the citric acid cycle (see Figure 2–82). Its two electrons are passed directly to ubiquinone, bypassing NADH dehydrogenase; they therefore cause less H^+ pumping than the two electrons transported from NADH.

We discuss next how the changes in redox potential from one electron carrier to the next are harnessed to pump protons out of the mitochondrial matrix.

A Large Drop in Redox Potential Across Each of the Three Respiratory Enzyme Complexes Provides the Energy for H^+ Pumping

We have previously discussed how the redox potential reflects electron affinities (see p. 76). **Figure 14–29** presents an outline of the redox potentials measured along the respiratory chain. These potentials drop in three large steps, one across each major respiratory complex. The change in redox potential between any two electron carriers is directly proportional to the free energy released when an electron transfers between them. Each enzyme complex acts as an energy-conversion device by harnessing some of this free-energy change to pump H^+ across the inner membrane, thereby creating an electrochemical proton gradient as electrons pass through that complex. This conversion can be demonstrated by purifying each respiratory enzyme complex and incorporating it separately into liposomes: when an appropriate electron donor and acceptor are added so that electrons can pass through the complex, protons are translocated across the liposome membrane.

The H^+ Pumping Occurs by Distinct Mechanisms in the Three Major Enzyme Complexes

Some respiratory enzyme complexes pump one H^+ per electron across the inner mitochondrial membrane, whereas others pump two. The detailed mechanism by which electron transport is coupled to H^+ pumping is different for the three different enzyme complexes. In the cytochrome b-c_1 complex, the quinones clearly have a role. As mentioned previously, a quinone picks up a H^+ from the aqueous medium along with each electron it carries and liberates it when it releases the electron (see Figure 14–24). Since ubiquinone is freely mobile in the lipid bilayer, it could accept electrons near the inside surface of the membrane and donate them to the cytochrome b-c_1 complex near the outside surface, thereby transferring one H^+ across the bilayer for every electron transported. Two protons are pumped per electron in the cytochrome b-c_1 complex, however. The complicated series of electron transfers that make this possible are still being worked out at the atomic level, aided by the complete structure of the cytochrome b-c_1 complex determined by x-ray crystallography (**Figure 14–30**).

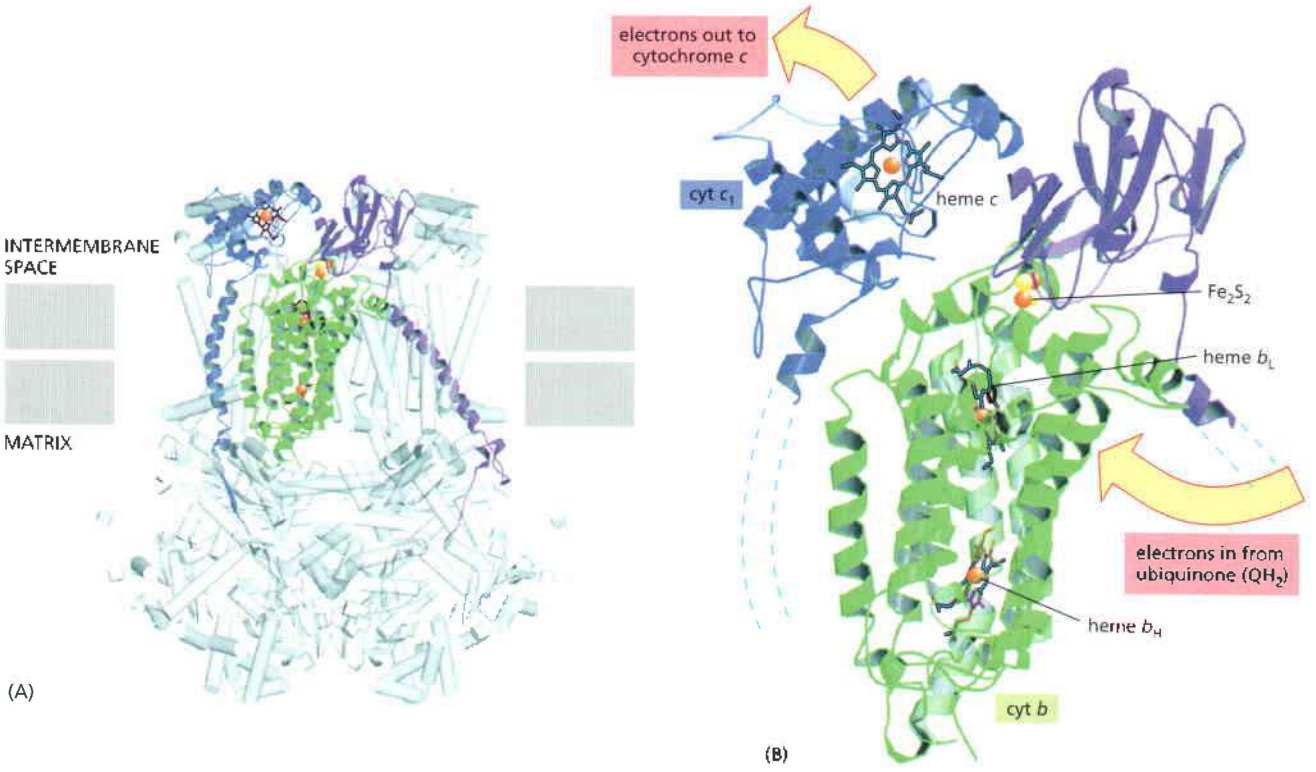


Figure 14–30 The atomic structure of cytochrome *b*-c_1. This protein is a dimer. The 240,000-dalton monomer is composed of 11 different protein molecules in mammals. The three colored proteins form the functional core of the enzyme: cytochrome *b (green)*, cytochrome c_1 *(blue)*, and the Rieske protein containing an iron–sulfur center *(purple)*. (A) The interaction of these three proteins across the two monomers. (B) Their electron carriers, along with the entrance and exit sites for electrons. The electrons initially donated by ubiquinone follow a complex path of electron and proton transfer reactions through the protein complex that enhances the storage of redox energy. This process, in which some of the electrons are recycled back into the quinone pool, is known as the *Q cycle*.

Electron transport causes allosteric changes in protein conformations that can also pump H^+, just as H^+ is pumped when ATP is hydrolyzed by the ATP synthase running in reverse. For both the NADH dehydrogenase complex and the cytochrome oxidase complex, it seems likely that electron transport drives sequential allosteric changes in protein conformation by altering the redox state of the components. These conformational changes in turn cause the protein to pump H^+ across the mitochondrial inner membrane. This type of H^+ pumping requires at least three distinct conformations for the pump protein; a general mechanism is presented in **Figure 14–31**.

Now that we have discussed the mechanistic basis for electron transport and proton pumping, we are ready to consider how the respiratory chain is regulated to make it optimally useful to the cell.

H^+ Ionophores Uncouple Electron Transport from ATP Synthesis

Since the 1940s, several substances—such as 2,4-dinitrophenol—have been known to act as *uncoupling agents*, uncoupling electron transport from ATP synthesis. The addition of these low-molecular-weight organic compounds to cells stops ATP synthesis by mitochondria without blocking their uptake of oxygen. In the presence of an uncoupling agent, electron transport and H^+ pumping continue at a rapid rate, but no H^+ gradient is generated. The explanation for this effect is both simple and elegant: uncoupling agents are lipid-soluble weak acids that act as diffusible H^+ carriers in the lipid bilayer (H^+ ionophores), and they provide a pathway for the flow of H^+ across the inner mitochondrial membrane that bypasses the ATP synthase. As a result of this short-circuiting, the proton-motive force is dissipated completely, and ATP can no longer be made.

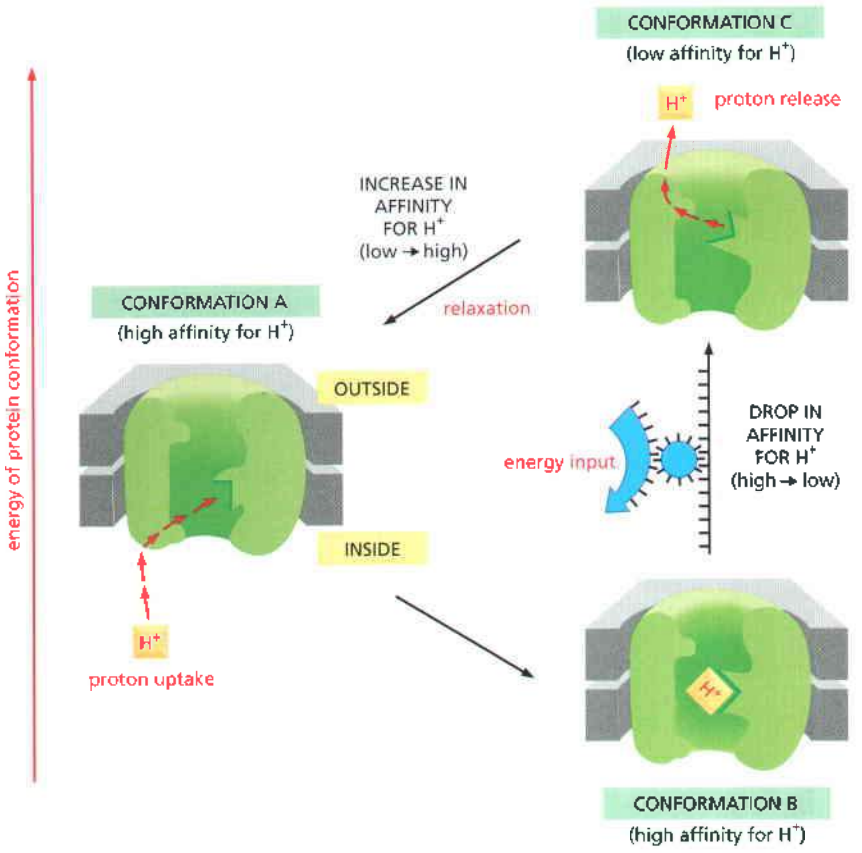


Figure 14–31 A general model for H^+ pumping. This model for H^+ pumping by a transmembrane protein is based on mechanisms that are thought to be used by NADH dehydrogenase and cytochrome oxidase, as well as by the light-driven procaryotic proton pump, bacteriorhodopsin. The protein is driven through a cycle of three conformations: A, B, and C. As indicated by their vertical spacing, these protein conformations have different energies. In conformation A, the protein has a high affinity for H^+, causing it to pick up a H^+ on the inside of the membrane. In conformation C, the protein has a low affinity for H^+, causing it to release a H^+ on the outside of the membrane. The transition from conformation B to conformation C that releases the H^+ is energetically unfavorable, and it occurs only because it is driven by being allosterically coupled to an energetically favorable reaction occurring elsewhere on the protein *(blue arrow)*. The other two conformational changes, A → B and C → A, lead to states of lower energy, and they proceed spontaneously. Because the overall cycle A → B → C → A releases free energy, H^+ is pumped from the inside (the matrix in mitochondria) to the outside (the intermembrane space in mitochondria). For NADH dehydrogenase and cytochrome oxidase, the energy required for the transition B → C is provided by electron transport, whereas for bacteriorhodopsin this energy is provided by light (see Figure 10–33). For yet other proton pumps, the energy is derived from ATP hydrolysis. In all cases, at least three distinct conformations are required to create a vectorial pumping process, for the same reason that at least three conformations are required to create a protein that can walk in a single direction along a filament (see Figure 3–77).

Respiratory Control Normally Restrains Electron Flow Through the Chain

The addition of an uncoupler such as dinitrophenol to cells causes mitochondria to increase their rate of electron transport substantially, resulting in an increase in oxygen uptake that reflects the existence of **respiratory control.** The control is thought to act via a direct inhibitory influence of the electrochemical proton gradient on the rate of electron transport. When an uncoupler collapses the gradient, electron transport is free to run unchecked at the maximal rate. As the gradient increases, electron transport becomes more difficult, and the process slows. Moreover, if an artificially large electrochemical proton gradient is experimentally created across the inner membrane, normal electron transport stops completely, and a *reverse electron flow* can be detected in some sections of the respiratory chain. This observation suggests that respiratory control reflects a simple balance between the free-energy change for electron-transport-linked proton pumping and the free-energy change for electron transport—that is, the magnitude of the electrochemical proton gradient affects both the rate and the direction of electron transport, just as it affects the directionality of the ATP synthase (see Figure 14–19).

Respiratory control is just one part of an elaborate interlocking system of feedback controls that coordinate the rates of glycolysis, fatty acid breakdown, the citric acid cycle, and electron transport. The rates of all of these processes are adjusted to the ATP:ADP ratio, increasing whenever an increased utilization of ATP causes the ratio to fall. The ATP synthase in the inner mitochondrial membrane, for example, works faster as the concentrations of its substrates ADP and P_i increase. As it speeds up, the enzyme lets more H^+ flow into the matrix and thereby dissipates the electrochemical proton gradient more rapidly. The falling gradient, in turn, enhances the rate of electron transport.

Similar controls, including feedback inhibition of several key enzymes by ATP, act to adjust the rates of NADH production to the rate of NADH utilization

by the respiratory chain, and so on. As a result of these multiple control mechanisms, the body oxidizes fats and sugars 5–10 times more rapidly during a period of strenuous exercise than during a period of rest.

Natural Uncouplers Convert the Mitochondria in Brown Fat into Heat-Generating Machines

In some specialized fat cells, mitochondrial respiration is normally uncoupled from ATP synthesis. In these cells, known as brown fat cells, most of the energy of oxidation is dissipated as heat rather than being converted into ATP. The inner membranes of the large mitochondria in these cells contain a special transport protein, known as an uncoupling protein, that allows protons to move down their electrochemical gradient without passing through ATP synthase. This uncoupling protein is switched on when heat generation is required, causing the cells to oxidize their fat stores at a rapid rate and produce more heat than ATP. Tissues containing brown fat serve as "heating pads," helping to revive hibernating animals and to protect sensitive areas of newborn human babies from the cold.

The Mitochondrion Has Many Critical Roles in Cell Metabolism

Cells are largely composed of macromolecules, which are constantly in need of repair or replacement as the cell ages. Even for cells and organisms that are not growing, those molecules that decay must be replaced by biosyntheses. Throughout this chapter, we emphasize the critical role of mitochondria in producing the ATP that cells need to maintain themselves as highly organized entities in a universe that is always driving toward increasing disorder (discussed in Chapter 2). In addition to ATP, however, the biosynthesis in the cytosol requires a constant supply of reducing power in the form of NADPH and of carbon skeletons. Most descriptions of this type of biosynthesis state that the needed carbon skeletons come directly from the breakdown of sugars, whereas the NADPH is produced in the cytosol by a side pathway for the breakdown of sugars (the pentose phosphate pathway, an alternative to glycolysis). But under conditions where foodstuffs are abundant, and plenty of ATP is available, mitochondria also generate both carbon skeletons and NADPH needed for cell growth. For this purpose, excess citrate produced in the mitochondrial matrix by the citric acid cycle is transported down its electrochemical gradient to the cytosol, where it is metabolized to produce both NADPH and carbon skeletons for biosyntheses. Thus, for example, as part of a cell's response to growth signals, large amounts of acetyl CoA are produced in the cytosol from citrate exported from mitochondria, accelerating the production of the fatty acids and sterols that build new membranes.

Mitochondria are also critical for buffering the redox potential in the cytosol. Cells need a constant supply of the electron acceptor NAD^+ for the central reaction in glycolysis that converts glyceraldehyde 3-phosphate to 1,3-bisphosphoglycerate (see Figure 2–72). This NAD^+ is converted to NADH in the process, and the NAD^+ needs to be regenerated by transferring the high-energy NADH electrons somewhere.

The NADH electrons will eventually be used to help drive oxidative phosphorylation inside the mitochondrion. But the mitochondrial inner membrane is impermeable to NADH. The electrons are therefore passed from the NADH to smaller molecules in the cytosol that can move through the inner mitochondrial membrane. Once in the matrix, these smaller molecules transfer their electrons to NAD^+ to form mitochondrial NADH, after which they are returned to the cytosol for recharging. This so-called shuttle system is bypassed in some specialized cells, such as insect flight muscle, that produce especially large amounts of ATP by aerobic glycolysis. Here the high-energy electrons derived from glyceraldehyde 3-phosphate are passed directly to the outer surface of the mitochondrial inner membrane, enabling them to enter the electron transport chain more rapidly and directly—but with loss of some of the usable energy.

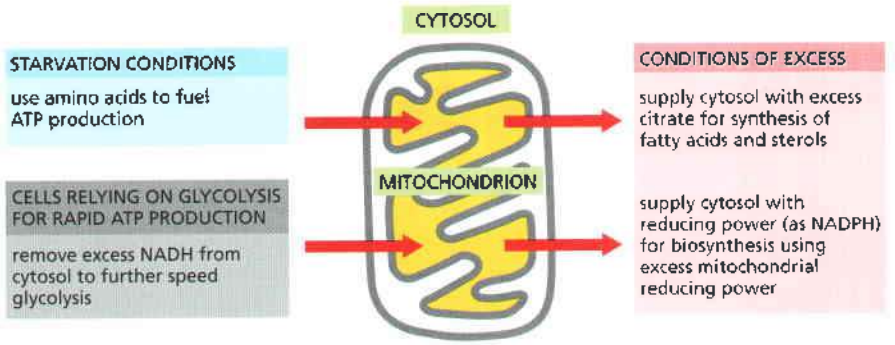


Figure 14–32 Critical roles of mitochondria in cell metabolism besides ATP production. The many essential metabolic reactions carried out by mitochondria, such as those illustrated here, emphasize the inadequacy of our emphasis on mitochondria as the cell furnace that oxidizes pyruvate and fatty acids to feed oxidative phosphorylation.

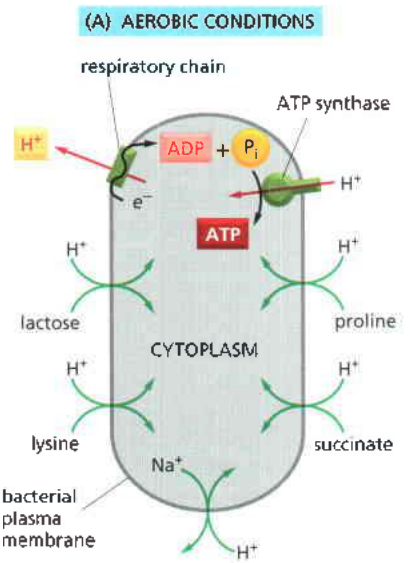
Under conditions of starvation, proteins in our bodies are broken down to amino acids, and the amino acids are imported into mitochondria and oxidized to produce NADH for ATP production. Thus, by carrying out different reactions under different conditions, the mitochondrion has many critical functions in maintaining cellular metabolism (Figure 14–32).

Bacteria Also Exploit Chemiosmotic Mechanisms to Harness Energy

Bacteria use enormously diverse energy sources. Some, like animal cells, are aerobic; they synthesize ATP from sugars they oxidize to CO_2 and H_2O by glycolysis, the citric acid cycle, and a respiratory chain in their plasma membrane that is similar to the one in the inner mitochondrial membrane. Others are strict anaerobes, deriving their energy either from glycolysis alone (by fermentation) or from an electron-transport chain that employs a molecule other than oxygen as the final electron acceptor. The alternative electron acceptor can be a nitrogen compound (nitrate or nitrite), a sulfur compound (sulfate or sulfite), or a carbon compound (fumarate or carbonate), for example. A series of electron carriers in the plasma membrane that are comparable to those in mitochondrial respiratory chains transfer the electrons to these acceptors.

Despite this diversity, the plasma membrane of the vast majority of bacteria contains an ATP synthase that is very similar to the one in mitochondria. In bacteria that use an electron-transport chain to harvest energy, the electron-transport chain pumps H^+ out of the cell and thereby establishes a proton-motive force across the plasma membrane that drives the ATP synthase to make ATP. In other bacteria, the ATP synthase works in reverse, using the ATP produced by glycolysis to pump H^+ and establish a proton gradient across the plasma membrane. The ATP used for this process is generated by fermentation processes (discussed in Chapter 2).

Thus, most bacteria, including the strict anaerobes, maintain a proton gradient across their plasma membrane. It can be harnessed to drive a flagellar motor, and it is used to pump Na^+ out of the bacterium via a Na^+-H^+ antiporter that takes the place of the Na^+-K^+ pump of eucaryotic cells. This gradient is also used for the active inward transport of nutrients, such as most amino acids and many sugars: each nutrient is dragged into the cell along with one or more protons through a specific symporter (Figure 14–33). In animal cells, by contrast,



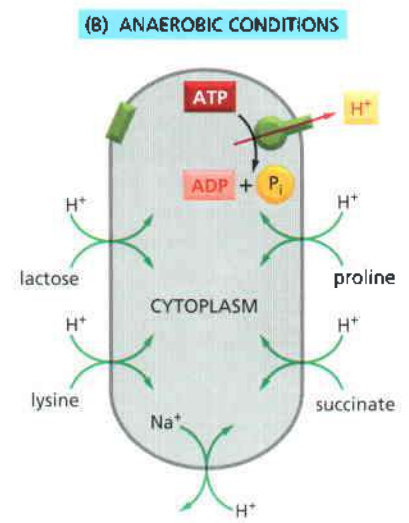


Figure 14–33 The importance of H^+-driven transport in bacteria. A proton-motive force generated across the plasma membrane pumps nutrients into the cell and expels Na^+. (A) In an aerobic bacterium, a respiratory chain produces an electrochemical proton gradient across the plasma membrane; this gradient is then used both to transport some nutrients into the cell and to make ATP. (B) The same bacterium growing under anaerobic conditions derives its ATP from glycolysis. The ATP synthase then hydrolyzes some of this ATP to establish an electrochemical proton gradient that drives those transport processes that depend on the respiratory chain in (A).

most inward transport across the plasma membrane is driven by the Na^+ gradient (high Na^+ outside, low Na^+ inside) that is established by the Na^+-K^+ pump.

Some unusual bacteria have adapted to live in a very alkaline environment and yet must maintain their cytoplasm at a physiological pH. For these cells, any attempt to generate an electrochemical H^+ gradient would be opposed by a large H^+ concentration gradient in the wrong direction (H^+ higher inside than outside). Presumably for this reason, some of these bacteria substitute Na^+ for H^+ in all of their chemiosmotic mechanisms. The respiratory chain pumps Na^+ out of the cell, the transport systems and flagellar motor are driven by an inward flux of Na^+, and a Na^+-driven ATP synthase synthesizes ATP. The existence of such bacteria demonstrates that the principle of chemiosmosis is more fundamental than the proton-motive force on which it is normally based.

Summary

The respiratory chain embedded in the inner mitochondrial membrane contains three respiratory enzyme complexes through which electrons pass on their way from NADH to O_2. Each complex can be purified, inserted into synthetic lipid vesicles, and then shown to pump H^+ when electrons are transported through it. In these complexes, electrons are transferred along a series of protein-bound electron carriers, including hemes and iron-sulfur centers. Energy released as the electrons move to lower and lower energy levels is used to drive allosteric changes in each respiratory enzyme complex that help to pump protons. Electrons are carried between enzyme complexes by the mobile electron carriers ubiquinone and cytochrome c to complete the electron-transport chain. The path of electron flow is NADH → NADH dehydrogenase complex → ubiquinone → cytochrome b-c_1 complex → cytochrome c → cytochrome oxidase complex → molecular oxygen (O_2).

The coupling of the energetically favorable transport of electrons to the pumping of H^+ out of the matrix creates an electrochemical proton gradient. This gradient is harnessed to make ATP by the ATP synthase, through which H^+ flows back into the matrix. The universal presence of ATP synthase in mitochondria, chloroplasts, and procaryotes testifies to the central importance of chemiosmotic mechanisms in cells.

CHLOROPLASTS AND PHOTOSYNTHESIS

All animals and most microorganisms rely on the continual uptake of large amounts of organic compounds from their environment. These compounds provide both the carbon skeletons for biosynthesis and the metabolic energy that drives cellular processes. It is likely that the first organisms on the primitive Earth had access to an abundance of the organic compounds produced by geochemical processes, but it is clear that most of these original compounds were used up billions of years ago. Since that time, nearly all of the organic materials required by living cells have been produced by *photosynthetic organisms*, including many types of photosynthetic bacteria.

The most advanced photosynthetic bacteria are the cyanobacteria, which have minimal nutrient requirements. They use electrons from water and the energy of sunlight to convert atmospheric CO_2 into organic compounds—a process called *carbon fixation*. In the course of splitting water [in the overall reaction $nH_2O + nCO_2 \xrightarrow{\text{light}} (CH_2O)_n + nO_2$], they also liberate into the atmosphere the oxygen required for oxidative phosphorylation. As we see in this section, it is thought that the evolution of cyanobacteria from more primitive photosynthetic bacteria eventually made possible the development of abundant aerobic life forms.

In plants and algae, which developed much later, photosynthesis occurs in a specialized intracellular organelle—the **chloroplast**. Chloroplasts perform photosynthesis during the daylight hours. Photosynthetic cells use the immediate products of photosynthesis, NADPH and ATP, to produce many organic molecules. In plants, the products include a low-molecular-weight sugar (usually

sucrose) that these cells export to meet the metabolic needs of the many non-photosynthetic cells of the organism.

Biochemical and genetic evidence strongly suggests that chloroplasts are descended from oxygen-producing photosynthetic bacteria that were endocytosed and lived in symbiosis with primitive eucaryotic cells. Mitochondria are also generally believed to be descended from an endocytosed bacterium. The many differences between chloroplasts and mitochondria presumably reflect their different bacterial ancestors, as well as their subsequent evolutionary divergence. Nevertheless, the fundamental mechanisms involved in light-driven ATP synthesis in chloroplasts are similar to those that we have already discussed for respiration-driven ATP synthesis in mitochondria.

The Chloroplast Is One Member of the Plastid Family of Organelles

Chloroplasts are the most prominent members of the **plastid** family of organelles. Plastids are present in all living plant cells, each cell type having its own characteristic complement. All plastids share certain features. Most notably, all plastids in a particular plant species contain multiple copies of the same relatively small genome. In addition, an envelope composed of two concentric membranes encloses each plastid.

As discussed in Chapter 12 (see Figure 12–3), all plastids develop from *proplastids*, small organelles in the immature cells of plant meristems (**Figure 14–34A**). Proplastids develop according to the requirements of each differentiated cell, and the type that is present is determined in large part by the nuclear genome. If a leaf is grown in darkness, its proplastids enlarge and develop into *etioplasts*, which have a semicrystalline array of internal membranes containing a yellow chlorophyll precursor instead of chlorophyll. When exposed to light, the etioplasts rapidly develop into chloroplasts by converting this precursor to chlorophyll and by synthesizing new membrane pigments, photosynthetic enzymes, and components of the electron-transport chain.

Leucoplasts are plastids present in many epidermal and internal tissues that do not become green and photosynthetic. They are little more than enlarged proplastids. A common form of leucoplast is the *amyloplast* (Figure 14–34B), which accumulates the polysaccharide starch in storage tissues—a source of sugar for future use. In some plants, such as potatoes, the amyloplasts can grow to be as large as an average animal cell.

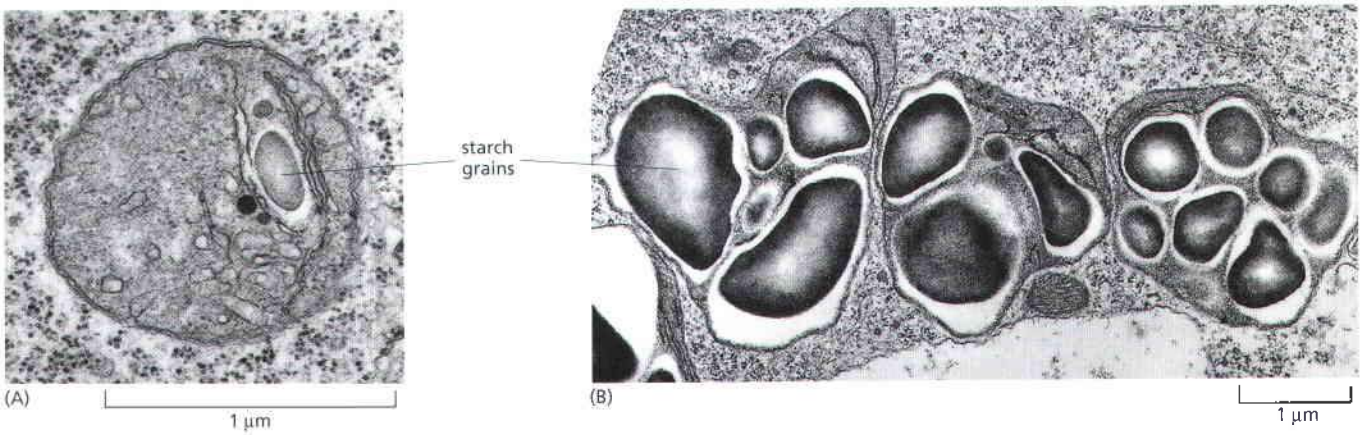
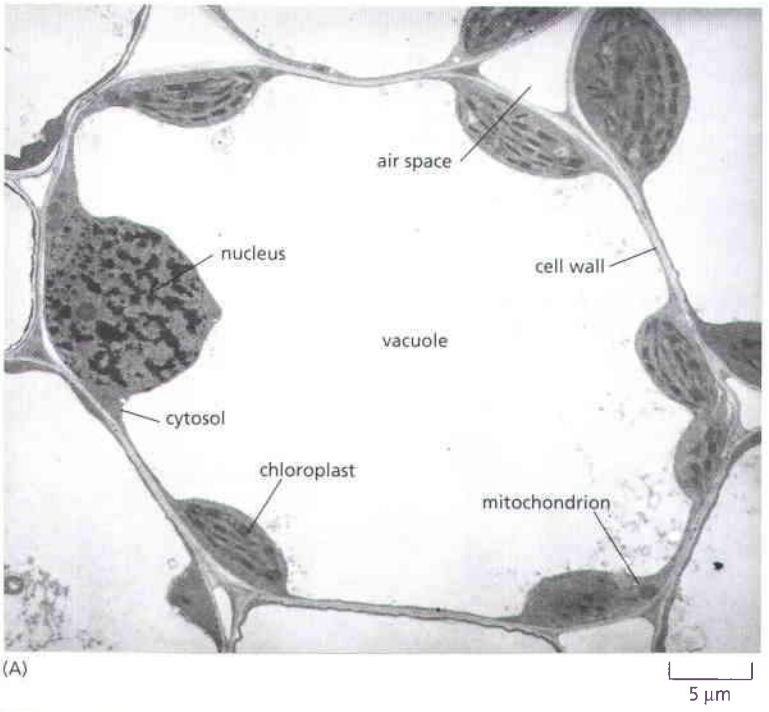
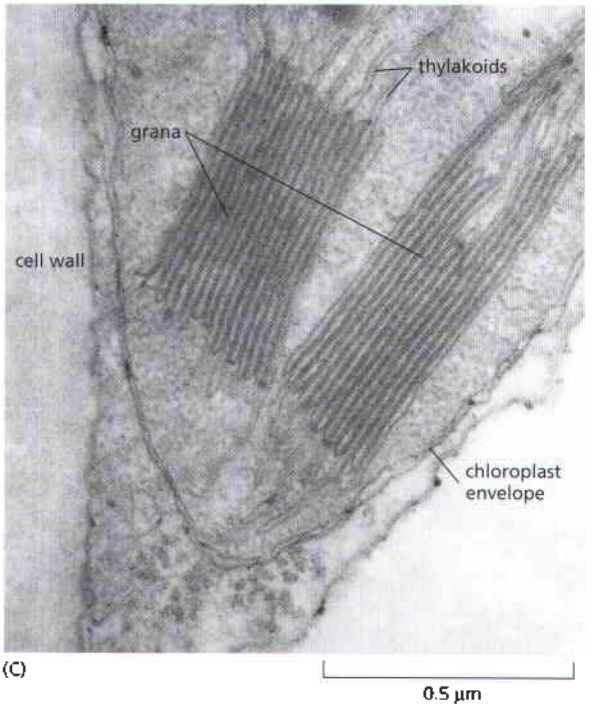


Figure 14–34 **Plastid diversity.** (A) A proplastid from a root tip cell of a bean plant. Note the double membrane; the inner membrane has also generated the relatively sparse internal membranes present. (B) Three amyloplasts (a form of leucoplast), or starch-storing plastids, in a root tip cell of soybean. (From B. Gunning and M. Steer, Plant Cell Biology: Structure and Function. Sudbury, MA: Jones & Bartlett, 1996.)





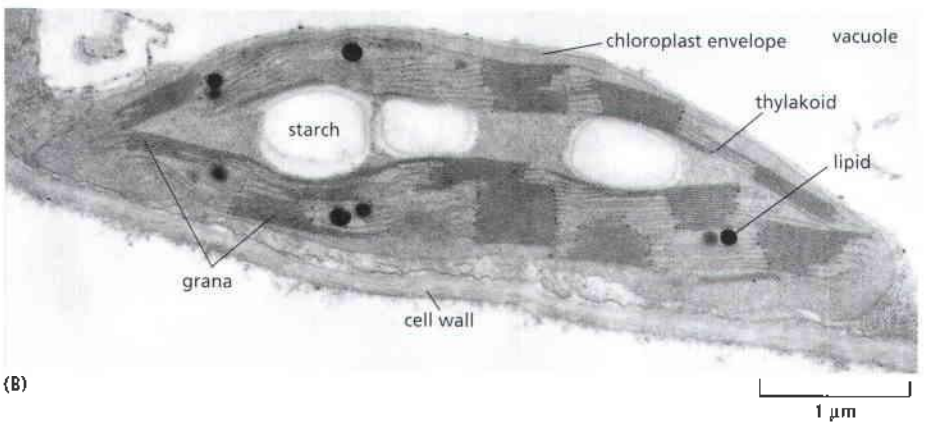


Figure 14–35 Electron micrographs of chloroplasts. (A) In a wheat leaf cell, a thin rim of cytoplasm—containing chloroplasts, the nucleus, and mitochondria—surrounds a large vacuole. (B) A thin section of a single chloroplast, showing the chloroplast envelope, starch granules, and lipid (fat) droplets that have accumulated in the stroma as a result of the biosyntheses occurring there. (C) A high-magnification view of two grana. A granum is a stack of thylakoids. (Courtesy of K. Plaskitt.)

Plastids are not just sites for photosynthesis and the deposition of storage materials. Plants have also used their plastids to compartmentalize their intermediary metabolism. Purine and pyrimidine synthesis, most amino acid synthesis, and all of the fatty acid synthesis of plants takes place in the plastids, whereas in animal cells these compounds are produced in the cytosol.

Chloroplasts Resemble Mitochondria But Have an Extra Compartment

Chloroplasts use chemiosmotic mechanisms to carry out their energy interconversions in much the same way that mitochondria do. Although much larger (**Figure 14–35**A), they are organized on the same principles. They have a highly permeable outer membrane; a much less permeable inner membrane, in which membrane transport proteins are embedded; and a narrow intermembrane space in between. Together, these membranes form the chloroplast envelope (Figure 14–35B,C). The inner membrane surrounds a large space called the **stroma**, which is analogous to the mitochondrial matrix and contains many metabolic enzymes. Like the mitochondrion, the chloroplast has its own genome and genetic system. The stroma therefore also contains a special set of ribosomes, RNAs, and the chloroplast DNA.

There is, however, an important difference between the organization of mitochondria and that of chloroplasts. The inner membrane of the chloroplast is not folded into cristae and does not contain electron-transport chains.

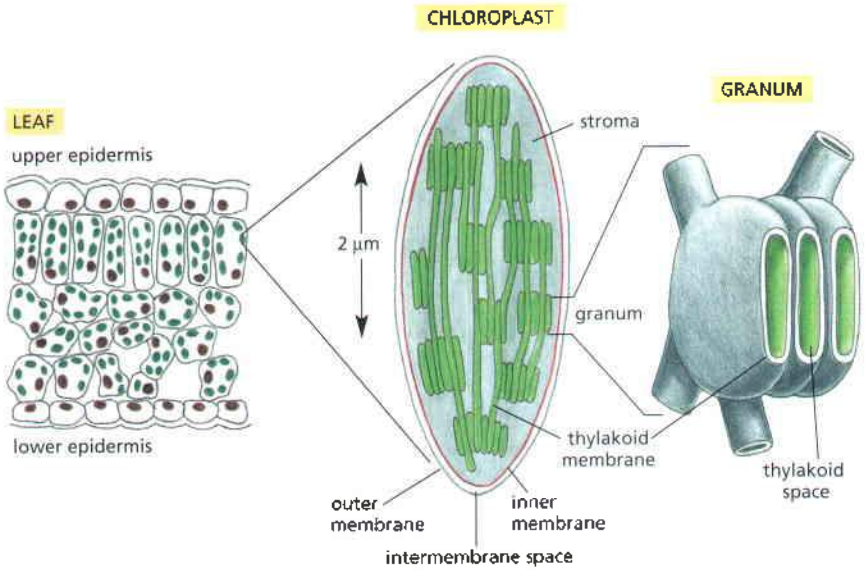


Figure 14–36 The chloroplast. This photosynthetic organelle contains three distinct membranes (the outer membrane, the inner membrane, and the thylakoid membrane) that define three separate internal compartments (the intermembrane space, the stroma, and the thylakoid space). The thylakoid membrane contains all the energy-generating systems of the chloroplast, including its chlorophyll. In electron micrographs, this membrane seems to be broken up into separate units that enclose individual flattened vesicles (see Figure 14–35), but these are probably joined into a single, highly folded membrane in each chloroplast. As indicated, the individual thylakoids are interconnected, and they tend to stack to form grana.

Instead, the electron-transport chains, photosynthetic light-capturing systems, and ATP synthase are all contained in the *thylakoid membrane*, a third distinct membrane that forms a set of flattened disclike sacs, the *thylakoids* (**Figure 14–36**). The lumen of each thylakoid is thought to be connected with the lumen of other thylakoids, thereby defining a third internal compartment called the *thylakoid space*, which is separated by the thylakoid membrane from the stroma that surrounds it. Thylakoid membranes interact with each other to form numerous local stacks called *grana*.

Figure 14–37 highlights the structural similarities and differences between mitochondria and chloroplasts. An important difference is that the head of the ATP synthase, where the ATP is made, protrudes into the stroma from the thylakoid membrane in a chloroplast, whereas it protrudes into the matrix from the inner mitochondrial membrane in a mitochondrion.

Chloroplasts Capture Energy from Sunlight and Use It to Fix Carbon

We can group the many reactions that occur during photosynthesis in plants into two broad categories:

1. In the **photosynthetic electron-transfer** reactions (also called the "light reactions"), energy derived from sunlight energizes an electron in the green organic pigment *chlorophyll*, enabling the electron to move along an electron-transport chain in the thylakoid membrane in much the same way that an electron moves along the respiratory chain in mitochondria.

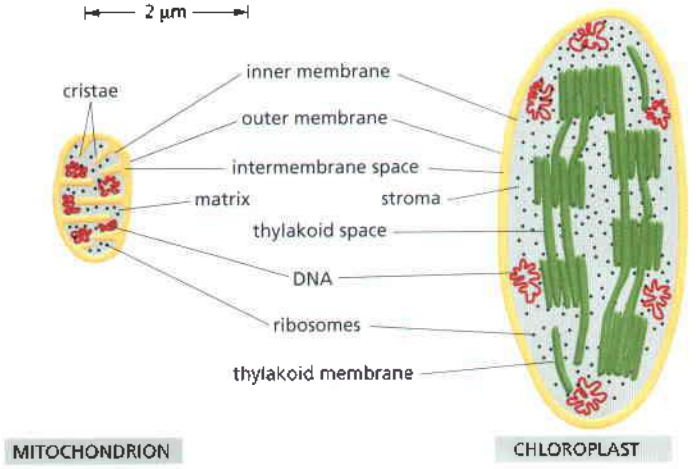


Figure 14–37 A mitochondrion and chloroplast compared. A chloroplast is generally much larger than a mitochondrion and contains, in addition to an outer and inner membrane, a thylakoid membrane enclosing a thylakoid space. Unlike the chloroplast inner membrane, the inner mitochondrial membrane is folded into cristae to increase its surface area.

The chlorophyll obtains its electrons from water (H_2O), producing O_2 as a by-product. During the electron-transport process, H^+ is pumped across the thylakoid membrane, and the resulting electrochemical proton gradient drives the synthesis of ATP in the stroma. As the final step in this series of reactions, high-energy electrons are loaded (together with H^+) onto $NADP^+$, converting it to NADPH. All of these reactions are confined to the chloroplast.

2. In the **carbon-fixation reactions** (also called the "dark reactions"), the ATP and the NADPH produced by the photosynthetic electron-transfer reactions serve as the source of energy and reducing power, respectively, to drive the conversion of CO_2 to carbohydrate. The carbon-fixation reactions, which begin in the chloroplast stroma and continue in the cytosol, produce sucrose and many other organic molecules in the leaves of the plant. The sucrose is exported to other tissues as a source of both organic molecules and energy for growth.

Thus, the formation of ATP, NADPH, and O_2 (which requires light energy directly) and the conversion of CO_2 to carbohydrate (which requires light energy only indirectly) are separate processes (Figure 14–38), although elaborate feedback mechanisms interconnect the two. Several of the chloroplast enzymes required for carbon fixation, for example, are inactivated in the dark and reactivated by light-stimulated electron-transport processes.

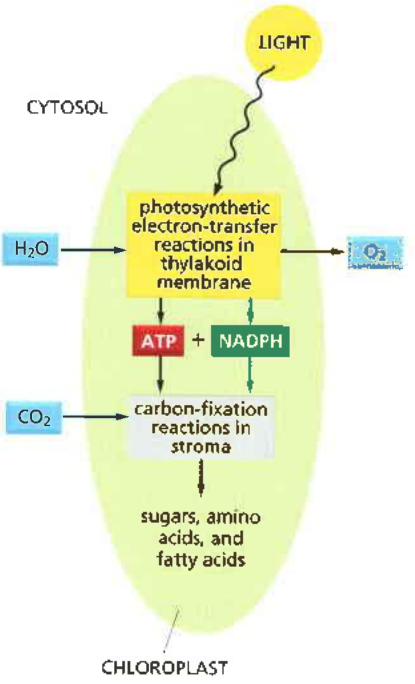


Figure 14–38 **The reactions of photosynthesis in a chloroplast.** Water is oxidized and oxygen is released in the photosynthetic electron-transfer reactions, while carbon dioxide is assimilated (fixed) to produce sugars and a variety of other organic molecules in the carbon-fixation reactions.

Carbon Fixation Is Catalyzed by Ribulose Bisphosphate Carboxylase

We have seen earlier in this chapter how cells produce ATP by using the large amount of free energy released when carbohydrates are oxidized to CO_2 and H_2O. Clearly, therefore, the reverse reaction, in which CO_2 and H_2O combine to make carbohydrate, must be a very unfavorable one that can occur only if it is coupled to other, very favorable reactions that drive it.

Figure 14–39 illustrates the central reaction of **carbon fixation**, in which an atom of inorganic carbon is converted to organic carbon: CO_2 from the atmosphere combines with the five-carbon compound ribulose 1,5-bisphosphate plus water to yield two molecules of the three-carbon compound 3-phosphoglycerate. This "carbon-fixing" reaction, which was discovered in 1948, is catalyzed in the chloroplast stroma by a large enzyme called *ribulose bisphosphate carboxylase*. Since each molecule of the complex works sluggishly (processing only about 3 molecules of substrate per second compared to 1000 molecules per second for a typical enzyme), the reaction requires an unusually large number of enzyme molecules. Ribulose bisphosphate carboxylase often constitutes more than 50% of the total chloroplast protein, and it is thought to be the most abundant protein on Earth.

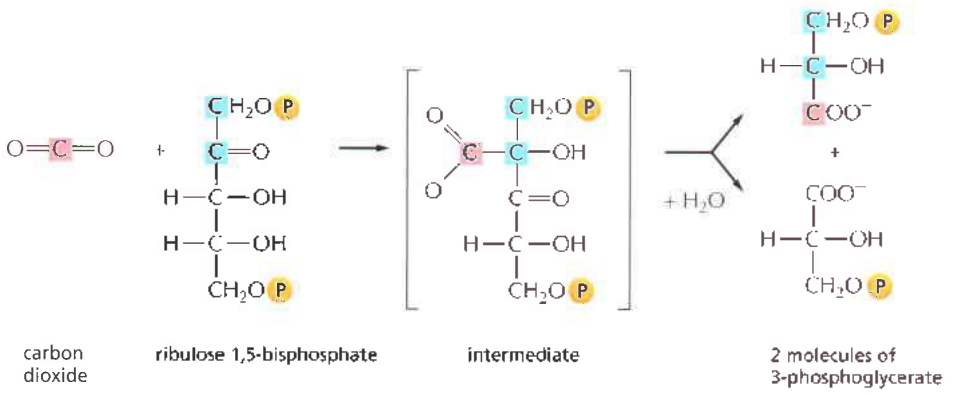


Figure 14–39 **The initial reaction in carbon fixation.** This reaction, in which carbon dioxide is converted into organic carbon, is catalyzed in the chloroplast stroma by the abundant enzyme ribulose bisphosphate carboxylase. The product is 3-phosphoglycerate, which is also an intermediate in glycolysis. The two carbon atoms shaded in *blue* are used to produce phosphoglycolate when the same enzyme adds oxygen instead of CO_2 (see text).

Each CO_2 Molecule That Is Fixed Consumes Three Molecules of ATP and Two Molecules of NADPH

The actual reaction in which CO_2 is fixed is energetically favorable because of the reactivity of the energy-rich compound *ribulose 1,5-bisphosphate*, to which each molecule of CO_2 is added (see Figure 14–39). The elaborate metabolic pathway that produces ribulose 1,5-bisphosphate requires both NADPH and ATP; it was worked out in one of the first successful applications of radioisotopes as tracers in biochemistry. This **carbon-fixation cycle** (also called the **Calvin cycle**) is outlined in **Figure 14–40**. It starts when 3 molecules of CO_2 are fixed by ribulose bisphosphate carboxylase to produce 6 molecules of 3-phosphoglycerate (containing $6 \times 3 = 18$ carbon atoms in all: 3 from the CO_2 and 15 from ribulose 1,5-bisphosphate). The 18 carbon atoms then undergo a cycle of reactions that regenerates the 3 molecules of ribulose 1,5-bisphosphate used in the initial carbon-fixation step (containing $3 \times 5 = 15$ carbon atoms). This leaves 1 molecule of *glyceraldehyde 3-phosphate* (3 carbon atoms) as the net gain.

Each CO_2 molecule converted into carbohydrate consumes a total of 3 molecules of ATP and 2 molecules of NADPH. The net equation is:

$$3CO_2 + 9ATP + 6NADPH + water \rightarrow$$
$$glyceraldehyde\ 3\text{-}phosphate + 8P_i + 9ADP + 6NADP^+$$

Thus, the formation of organic molecules from CO_2 and H_2O requires both *phosphate-bond energy* (as ATP) and *reducing power* (as NADPH). We return to this important point later.

The glyceraldehyde 3-phosphate produced in chloroplasts by the carbon-fixation cycle is a three-carbon sugar that also serves as a central intermediate in glycolysis. Much of it is exported to the cytosol, where it can be converted into

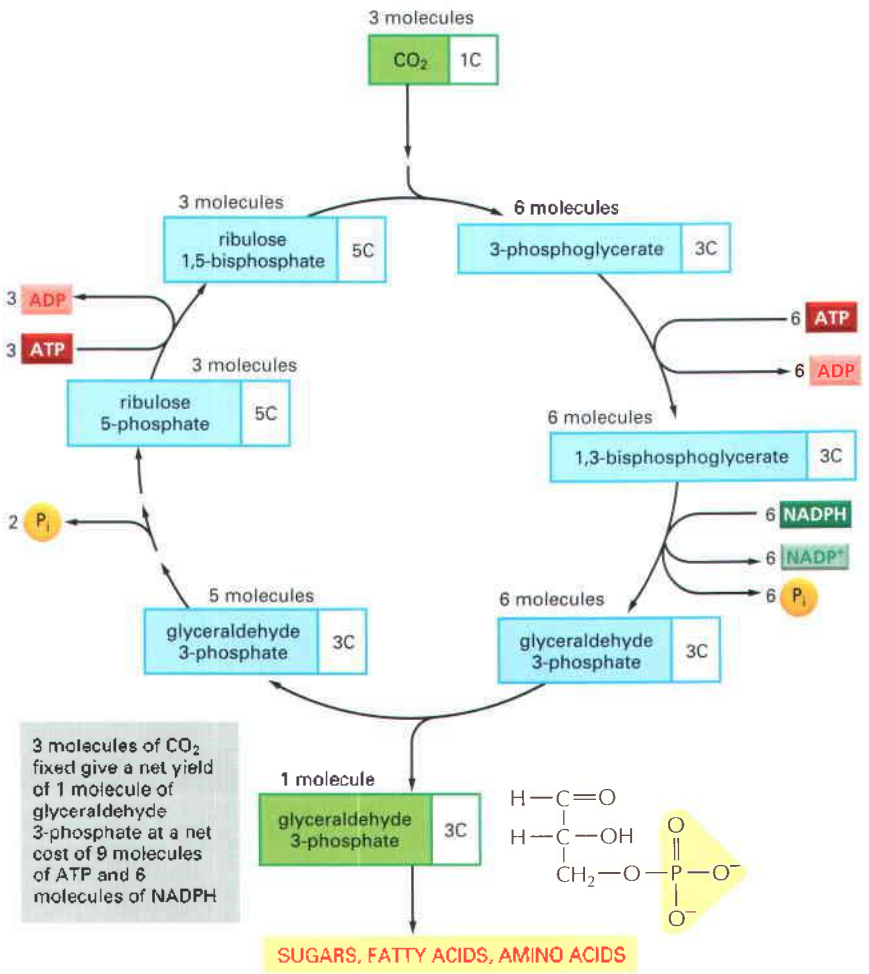


Figure 14–40 The carbon-fixation cycle, which forms organic molecules from CO_2 and H_2O. The number of carbon atoms in each type of molecule is indicated in the *white box*. There are many intermediates between glyceraldehyde 3-phosphate and ribulose 5-phosphate, but they have been omitted here for clarity. The entry of water into the cycle is also not shown.

fructose 6-phosphate and glucose 1-phosphate by the reversal of several reactions in glycolysis (see Panel 2–8, pp. 120–121). The glucose 1-phosphate is then converted to the sugar nucleotide UDP-glucose, and this combines with the fructose 6-phosphate to form sucrose phosphate, the immediate precursor of the disaccharide sucrose. **Sucrose** is the major form in which sugar is transported between plant cells: just as glucose is transported in the blood of animals, sucrose is exported from the leaves via vascular bundles, providing the carbohydrate required by the rest of the plant.

Most of the glyceraldehyde 3-phosphate that remains in the chloroplast is converted to *starch* in the stroma. Like glycogen in animal cells, **starch** is a large polymer of glucose that serves as a carbohydrate reserve (see Figure 14–34B). The production of starch is regulated so that it is produced and stored as large grains in the chloroplast stroma during periods of excess photosynthetic capacity. This occurs through reactions in the stroma that are the reverse of those in glycolysis: they convert glyceraldehyde 3-phosphate to glucose 1-phosphate, which is then used to produce the sugar nucleotide ADP-glucose, the immediate precursor of starch. At night the plant breaks down the starch to help support the metabolic needs of the plant. Starch provides an important part of the diet of all animals that eat plants.

Carbon Fixation in Some Plants Is Compartmentalized to Facilitate Growth at Low CO_2 Concentrations

Although ribulose bisphosphate carboxylase preferentially adds CO_2 to ribulose 1,5-bisphosphate, it can use O_2 as a substrate in place of CO_2, and if the concentration of CO_2 is low, it will add O_2 to ribulose 1,5-bisphosphate instead (see Figure 14–39). This is the first step in a pathway called **photorespiration**, whose ultimate effect is to use up O_2 and liberate CO_2 without the production of useful energy stores. In many plants, about one-third of the CO_2 fixed is lost again as CO_2 because of photorespiration.

Photorespiration can be a serious liability for plants in hot, dry conditions, which cause them to close their stomata (the gas exchange pores in their leaves, each of which is called a stoma) to avoid excessive water loss. This in turn causes the CO_2 levels in the leaf to fall precipitously, thereby favoring photorespiration. A special adaptation, however, occurs in the leaves of many plants, such as corn and sugar cane, that grow in hot, dry environments. In these plants, the carbon-fixation cycle occurs only in the chloroplasts of specialized *bundle-sheath cells*, which contain all of the plant's ribulose bisphosphate carboxylase. These cells are protected from the air and are surrounded by a specialized layer of *mesophyll cells* that use the energy harvested by their chloroplasts to "pump" CO_2 into the bundle-sheath cells. This supplies the ribulose bisphosphate carboxylase with a high concentration of CO_2, thereby greatly reducing photorespiration.

The CO_2 pump is produced by a reaction cycle that begins in the cytosol of the mesophyll cells. A CO_2-fixation step is catalyzed by an enzyme that binds carbon dioxide (as bicarbonate) and combines it with an activated three-carbon molecule (phosphoenol-pyruvate) to produce a four-carbon molecule. The four-carbon molecule diffuses into the bundle-sheath cells, where it is broken down to release the CO_2 and generate a molecule with three carbons. The pumping cycle is completed when this three-carbon molecule is returned to the mesophyll cells and converted back to its original activated form. Because the CO_2 is initially captured by converting it into a compound containing four carbons, the CO_2-pumping plants are called *C_4 plants*. All other plants are called *C_3 plants* because they capture CO_2 into the three-carbon compound 3-phosphoglycerate (**Figure 14–41**).

As with any vectorial transport process, pumping CO_2 into the bundle-sheath cells in C_4 plants costs energy (ATP is hydrolyzed; see Figure 14–41B). In hot, dry environments, however, this cost can be much less than the energy lost by photorespiration in C_3 plants, so C_4 plants have a potential advantage. Moreover, because C_4 plants can perform photosynthesis at a lower concentration of CO_2 inside the leaf, they need to open their stomata less often and therefore can

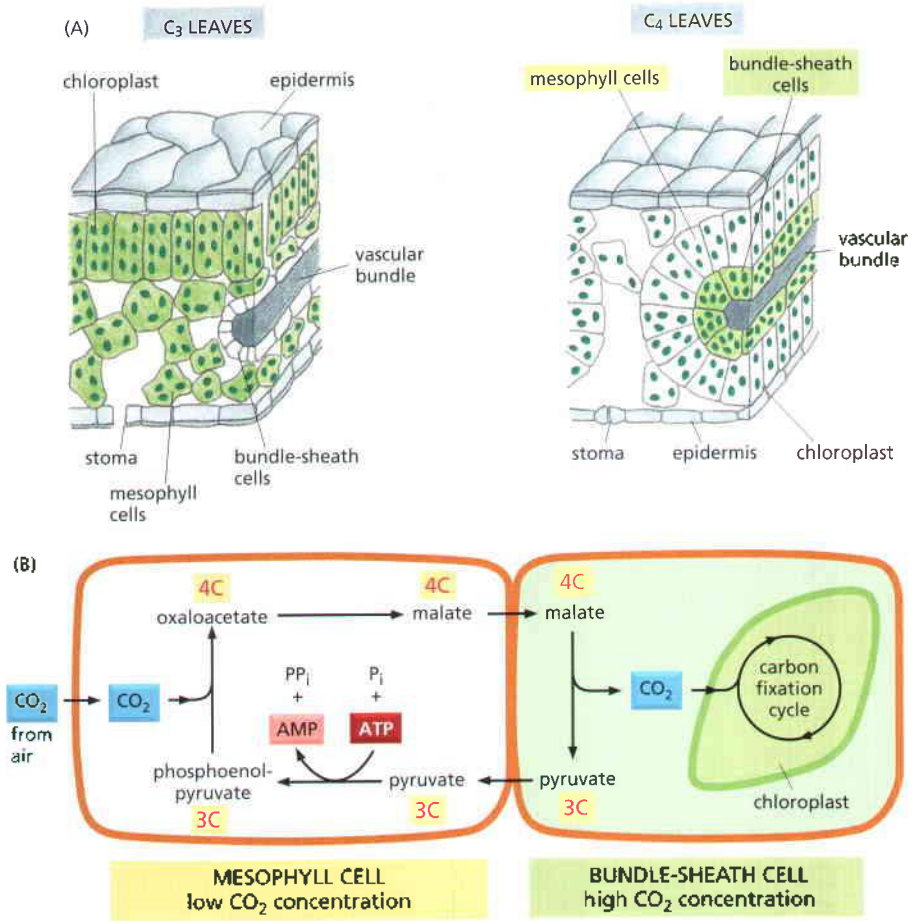


Figure 14–41 The CO_2 pumping in C_4 plants. (A) Comparative leaf anatomy in a C_3 plant and a C_4 plant. The cells with *green* cytosol in the leaf interior contain chloroplasts that perform the normal carbon-fixation cycle. In C_4 plants, the mesophyll cells are specialized for CO_2 pumping rather than for carbon fixation, and they thereby create a high ratio of CO_2 to O_2 in the bundle-sheath cells, which are the only cells in these plants where the carbon-fixation cycle occurs. The vascular bundles carry the sucrose made in the leaf to other tissues. (B) How carbon dioxide is concentrated in bundle-sheath cells by the harnessing of ATP energy in mesophyll cells.

fix about twice as much net carbon as C_3 plants per unit of water lost. This type of carbon fixation has evolved independently in several different plant lineages. Although the vast majority of plant species are C_3 plants, C_4 plants such as corn and sugar cane are much more effective at converting sunlight energy into biomass than C_3 plants such as cereal grains. They are therefore of special importance in world agriculture.

Photosynthesis Depends on the Photochemistry of Chlorophyll Molecules

Having discussed the carbon-fixation reactions, we now return to the question of how the photosynthetic electron-transfer reactions in the chloroplast generate the ATP and the NADPH needed to drive the production of carbohydrates from CO_2 and H_2O. Sunlight absorbed by **chlorophyll** molecules supplies the required energy (**Figure 14–42**). The process of energy conversion begins when a quantum of light (a photon) excites a chlorophyll molecule, causing an electron in the chlorophyll to move from one molecular orbital to another of higher energy. Such an excited molecule is unstable and tends to return quickly to its original, unexcited state. This can happen in one of three ways:

1. By converting the extra energy into heat (molecular motions) or to some combination of heat and light of a longer wavelength (fluorescence); this is what happens when an isolated chlorophyll molecule in solution absorbs light energy.
2. By transferring the energy—but not the electron—directly to a neighboring chlorophyll molecule by a process called *resonance energy transfer*.
3. By transferring the negatively charged high-energy electron to another nearby molecule, an *electron acceptor*, after which the positively charged chlorophyll returns to its original state by taking up a low-energy electron from some other molecule, an *electron donor*.

In the process of photosynthesis, the last two mechanisms are greatly facilitated by two different protein complexes: resonance energy transfer by an *antenna complex* and high-energy electron transfer by a *photochemical reaction center*. These two types of protein complexes, acting in concert, make most of the life on Earth possible. We shall now describe how they work.

A Photochemical Reaction Center Plus an Antenna Complex Form a Photosystem

Large multiprotein complexes called **photosystems** catalyze the conversion of the light energy captured in excited chlorophyll molecules to useful forms. A photosystem consists of two closely linked components: an antenna complex, consisting of proteins bound to a large set of pigment molecules that capture light energy and feed it to the reaction center; and a photochemical reaction center, consisting of a complex of proteins and chlorophyll molecules that enable light energy to be converted into chemical energy.

The **antenna complex** is important for capturing light energy. In chloroplasts it consists of a number of distinct membrane protein complexes (known as light-harvesting complexes); together, these proteins bind several hundred chlorophyll molecules per reaction center, orienting them precisely in the thylakoid membrane. The antenna complex also contains accessory pigments called *carotenoids*, which protect the chlorophylls from oxidation and can help collect light of other wavelengths. When light excites a chlorophyll molecule in the antenna complex, the energy is rapidly transferred from one molecule to another by resonance energy transfer until it reaches a special pair of chlorophyll molecules in the photochemical reaction center. Each antenna complex thereby acts as a funnel, collecting light energy and directing it to a specific site where it can be used effectively (**Figure 14–43**).

The **photochemical reaction center** is a transmembrane protein–pigment complex that lies at the heart of photosynthesis. It is thought to have evolved more than 3 billion years ago in primitive photosynthetic bacteria. The special pair of chlorophyll molecules in the reaction center acts as an irreversible trap for excitation quanta because the excited electron is immediately passed to a neighboring chain of electron acceptors in the protein complex (**Figure 14–44**). By moving the high-energy electron rapidly away from the chlorophylls, the photochemical reaction center transfers it to an environment where it is much more stable. The electron is thereby suitably positioned for subsequent reactions. These require more time to complete, and they result in the production of light-generated high-energy electrons that are fed into electron transport chains.

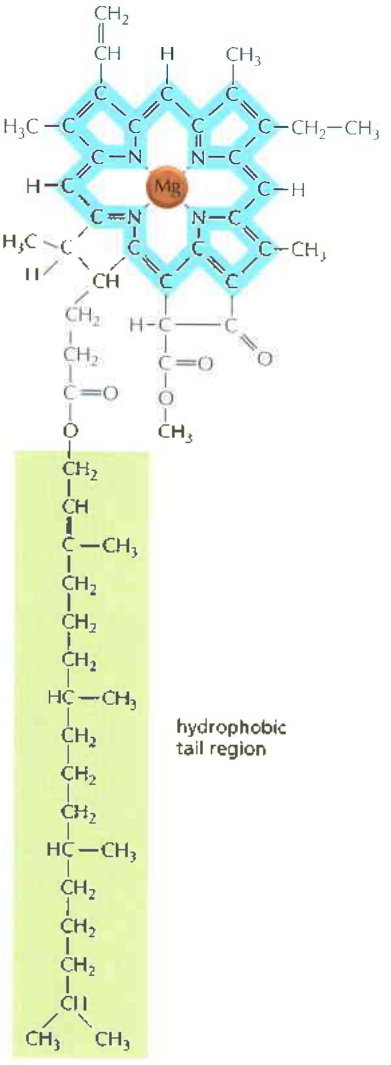


Figure 14–42 **The structure of chlorophyll.** A magnesium atom is held in a porphyrin ring, which is related to the porphyrin ring that binds iron in heme (see Figure 14–22). Electrons are delocalized over the bonds shown in *blue*.

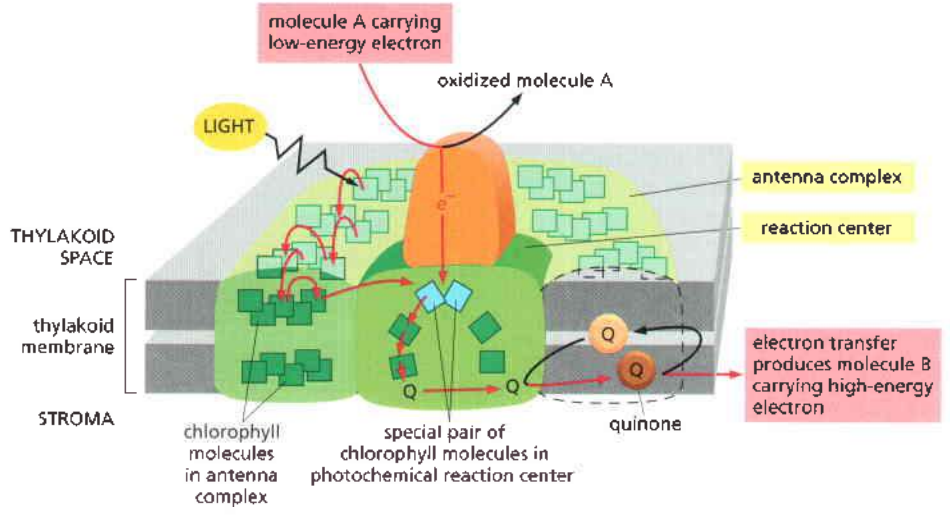


Figure 14–43 **The antenna complex and photochemical reaction center in a photosystem.** <GCAC> The antenna complex is a collector of light energy in the form of excited electrons. The energy of the excited electrons is funneled, through a series of resonance energy transfers, to a special pair of chlorophyll molecules in the photochemical reaction center. The reaction center then produces a high-energy electron that can be passed rapidly to the electron-transport chain in the thylakoid membrane, via a quinone.

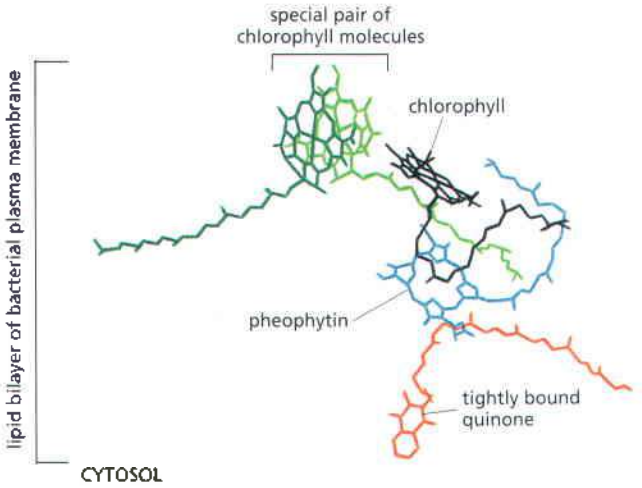


Figure 14–44 The arrangement of the electron carriers in the photochemical reaction center of a purple bacterium. <ATCA> The pigment molecules shown are held in the interior of a transmembrane protein and are surrounded by the lipid bilayer of the bacterial plasma membrane. An electron in the special pair of chlorophyll molecules is excited by resonance from an antenna complex chlorophyll, and the excited electron is then transferred stepwise from the special pair to the quinone (see also Figure 14–45). A similar arrangement of electron carriers is present in the reaction centers of plants (see Figure 14–47).

In a Reaction Center, Light Energy Captured by Chlorophyll Creates a Strong Electron Donor from a Weak One

The electron transfers involved in the photochemical reactions just described have been analyzed extensively by rapid spectroscopic methods. **Figure 14–45** illustrates, in a general way, how light provides the energy needed to transfer an electron from a weak electron donor (a molecule with a strong affinity for electrons) to a molecule that is a strong electron donor in its reduced form (a

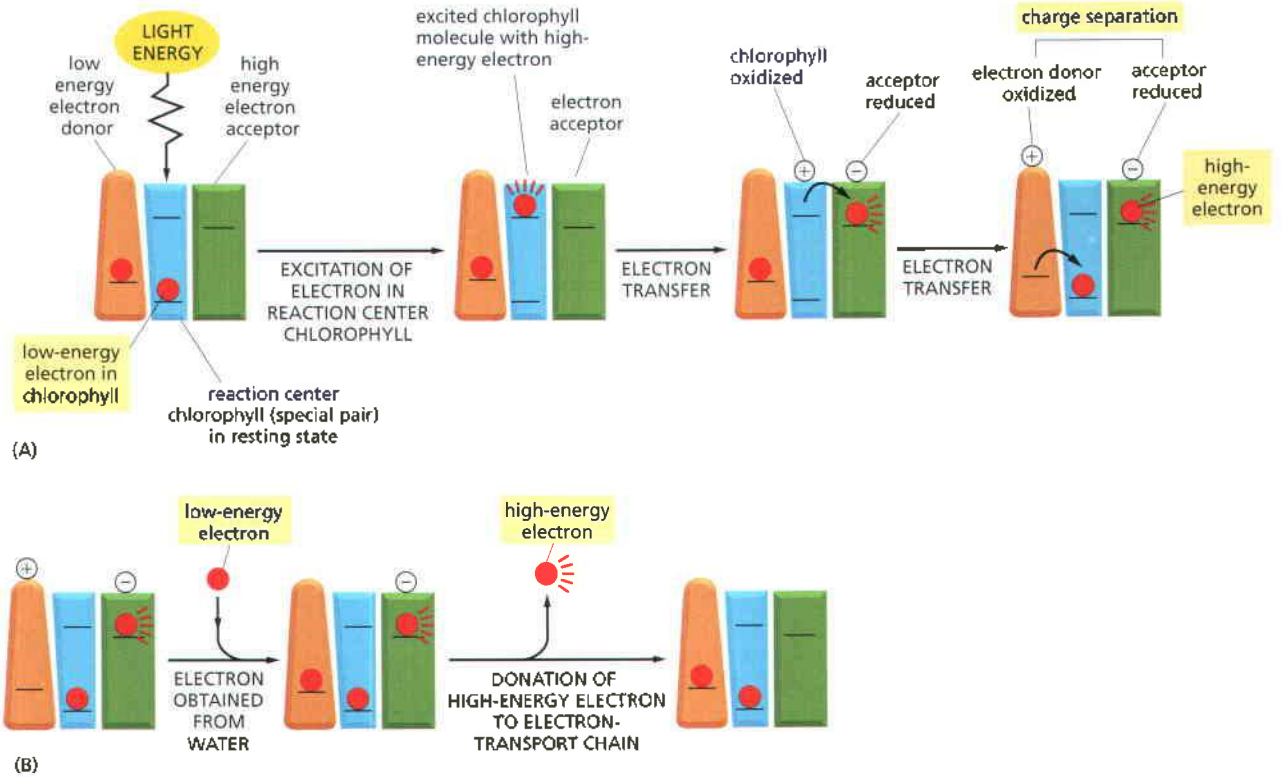


Figure 14–45 How light energy is harvested by a reaction center chlorophyll molecule. (A) The initial events in a reaction center create a charge separation. A pigment–protein complex holds a chlorophyll molecule of the special pair *(blue)* precisely positioned so that both a potential low-energy electron donor *(orange)* and a potential high-energy electron acceptor *(green)* are immediately available. When light energizes an electron in the chlorophyll molecule (*red* electron), the excited electron is immediately passed to the electron acceptor and is thereby partially stabilized. The positively charged chlorophyll molecule then quickly attracts the low-energy electron from the electron donor and returns to its resting state, creating a larger charge separation that further stabilizes the high-energy electron. These reactions require less than 10^{-6} second to complete. (B) In the final stage of this process, which follows the steps in (A), the photosynthetic reaction center is restored to its original resting state by acquiring a new low-energy electron and then transferring the high-energy electron derived from chlorophyll to an electron transport chain in the membrane. As will be discussed subsequently, the ultimate source of low-energy electrons for photosystem II in the chloroplast is water; as a result, light produces high-energy electrons in the thylakoid membrane from low-energy electrons in water.

molecule with a weak affinity for electrons). The special pair of chlorophyll molecules in the reaction center is poised to pass each excited electron to a precisely positioned neighboring molecule in the same protein complex (an electron acceptor). The chlorophyll molecule that loses an electron becomes positively charged, but it rapidly regains an electron from an adjacent electron donor to return to its unexcited, uncharged state (Figure 14–45A, *orange* electron). Then, in slower reactions, the electron donor has its missing electron replaced, and the high-energy electron that was generated by the excited chlorophyll is transferred to the electron-transport chain (Figure 14–45B). The excitation energy in chlorophyll that would normally be released as fluorescence or heat is thereby used instead to create a strong electron donor (a molecule carrying a high-energy electron) where none had been before.

The photosystem of purple bacteria is somewhat simpler than the evolutionarily related photosystems in chloroplasts, and it has served as a good model for working out reaction details. The reaction center in this photosystem is a large protein–pigment complex that can be solubilized with detergent and purified in active form. In a major triumph of structure analysis, its complete three-dimensional structure was determined by x-ray crystallography (see Figure 10–34). This structure, combined with kinetic data, provides the best picture we have of the initial electron-transfer reactions that occur during photosynthesis. **Figure 14–46** shows the actual sequence of electron transfers that take place, for comparison with Figure 14–45A.

In the purple bacterium, the electron used to fill the electron-deficient hole created by the light-induced charge separation comes from a cyclic flow of electrons transferred through a cytochrome (see *orange box* in Figure 14–45); the strong electron donor produced is a quinone. One of the two photosytems in the chloroplasts of higher plants likewise produces a quinone carrying high-energy electrons. However, as we discuss next, because water provides the electrons for this photosystem, photosynthesis in plants—unlike that in purple bacteria—releases large quantities of oxygen gas.

Noncyclic Photophosphorylation Produces Both NADPH and ATP

Photosynthesis in plants and cyanobacteria produces both ATP and NADPH directly by a two-step process called **noncyclic photophosphorylation**. Because two photosystems—called photosystems I and II—work in series to energize an electron to a high-enough energy state, the electron can be transferred all the

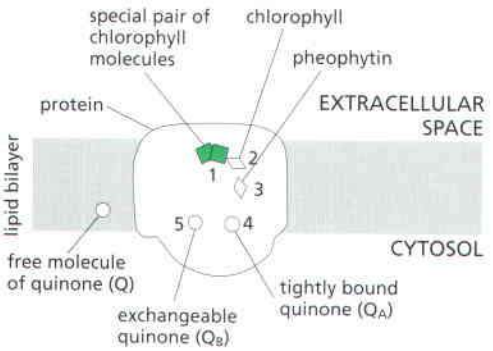
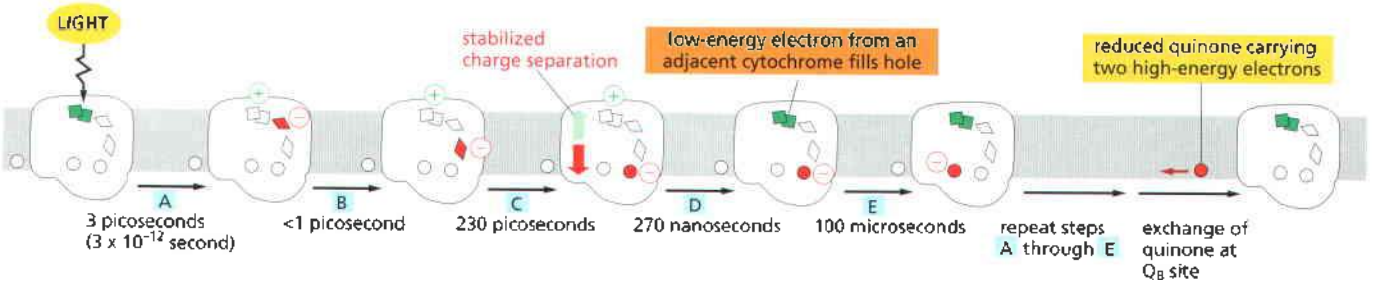


Figure 14–46 **The electron transfers that occur in the photochemical reaction center of a purple bacterium.** <ATCA> A similar set of reactions occurs in the evolutionarily related photosystem II in plants. At the top left is an orientating diagram showing the molecules that carry electrons, which are those in Figure 14–45, plus an exchangeable quinone (Q_B) and a freely mobile quinone (Q) dissolved in the lipid bilayer. Electron carriers 1–5 are each bound in a specific position on a 596-amino-acid transmembrane protein formed from two separate subunits (see Figure 10–34). After excitation by a photon of light, a high-energy electron passes from pigment molecule to pigment molecule, very rapidly creating a stable *charge separation*, as shown in the sequence of steps A–C, in which the pigment molecule carrying a high-energy electron is colored *red*. Steps D and E then occur progressively. After a second photon has repeated this sequence with a second electron, the exchangeable quinone is released into the bilayer carrying two high-energy electrons. This quinone quickly loses its charge by picking up two protons (see Figure 14–24).



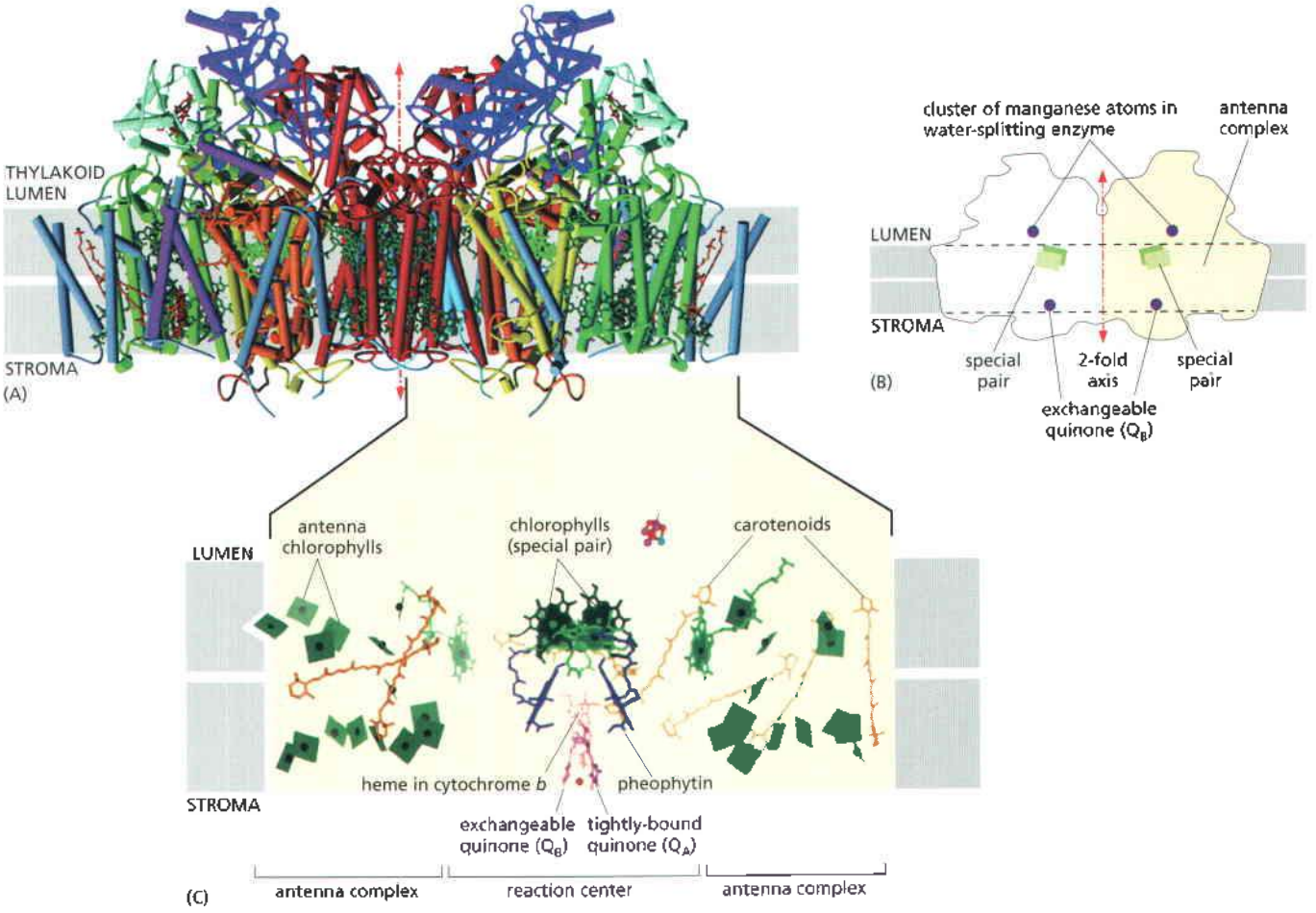


Figure 14–47 The structure of photosystem II in plants and cyanobacteria. The structure shown is a dimer, organized around a two fold axis (*red dotted* arrows). Each monomer is composed of 16 integral membrane protein subunits plus three subunits in the lumen, with a total of 36 bound chlorophylls, 7 carotenoids, two pheophytins, two hemes, two plastoquinones, and one manganese cluster in an oxygen-evolving water-splitting center. (A) The complete three-dimensional structure of the dimer. (B) Schematic of the dimer with a few central features indicated. (C) A monomer drawn to show only the non-protein molecules in the structure, thereby highlighting the protein-bound pigments and electron carriers; *green* structures are chlorophylls. (Adapted from K.N. Ferreira et al., *Science* 303:1831–1838, 2004. With permission from AAAS.)

way from water to NADPH. As the high-energy electrons pass through the coupled photosystems to generate NADPH, some of their energy is siphoned off for ATP synthesis.

The first of the two photosystems—paradoxically called *photosystem II* for historical reasons—has the unique ability to withdraw electrons from water. The oxygens of two water molecules bind to a cluster of four manganese atoms on the lumenal surface of the photosystem II reaction center complex (**Figure 14–47**). This cluster enables electrons to be removed one at a time from the water, as required to fill the electron-deficient holes created by light in chlorophyll molecules in the reaction center. As soon as four electrons have been removed from the two water molecules (requiring four quanta of light), O_2 is released. Photosystem II thus catalyzes the reaction $2H_2O + 4$ photons $\rightarrow 4H^+ + 4e^- + O_2$. As we discussed for the electron-transport chain in mitochondria, which uses O_2 and produces water, the mechanism ensures that no partly oxidized water molecules are released as dangerous, highly reactive oxygen radicals. Essentially all the oxygen in the Earth's atmosphere has been produced in this way.

The core of the reaction center in photosystem II is homologous to the bacterial reaction center just described, and it likewise produces strong electron donors in the form of reduced quinone molecules dissolved in the lipid bilayer of the membrane. The quinones pass their electrons to a H^+ pump called the *cytochrome b_6-f complex*, which resembles the cytochrome b-c_1 complex in the respiratory chain of mitochondria. The cytochrome b_6-f complex pumps H^+ into the thylakoid space across the thylakoid membrane (or out of the cytosol across the plasma membrane in cyanobacteria), and the resulting electrochemical gradient drives the synthesis of ATP by an ATP synthase (**Figure 14–48**).

The final electron acceptor in this electron-transport chain is the second photosystem, *photosystem I*, which accepts an electron into the electron-deficient hole created by light in the chlorophyll molecule in its reaction center. Each electron that enters photosystem I is finally boosted to a very high-energy

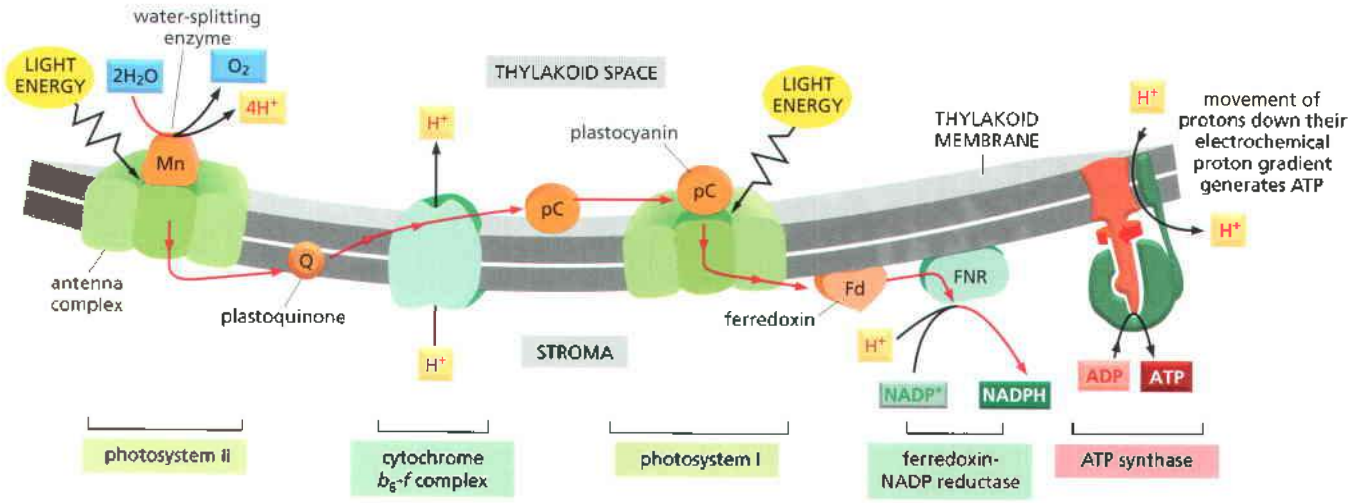


Figure 14–48 Electron flow during photosynthesis in the thylakoid membrane. The mobile electron carriers in the chain are plastoquinone (which closely resembles the ubiquinone of mitochondria), plastocyanin (a small copper-containing protein), and ferredoxin (a small protein containing an iron–sulfur center). The cytochrome b_6-f complex resembles the b-c_1 complex of mitochondria and the b-c complex of bacteria (see Figure 14–73): all three complexes accept electrons from quinones and pump H^+ across the membrane. The H^+ released by water oxidation into the thylakoid space, and the H^+ consumed during NADPH formation in the stroma, also contribute to the generation of the electrochemical H^+ gradient. As illustrated, this gradient drives ATP synthesis by an ATP synthase present in this same membrane.

level that allows it to be passed to the iron–sulfur center in ferredoxin and then to $NADP^+$ to generate NADPH.

The scheme for photosynthesis just discussed is known as the *Z scheme*. By means of its two electron-energizing steps, one catalyzed by each photosystem, an electron is passed from water, which normally holds on to its electrons very tightly (redox potential = +820 mV), to NADPH, which normally holds on to its electrons loosely (redox potential = –320 mV) (Figure 14–49).

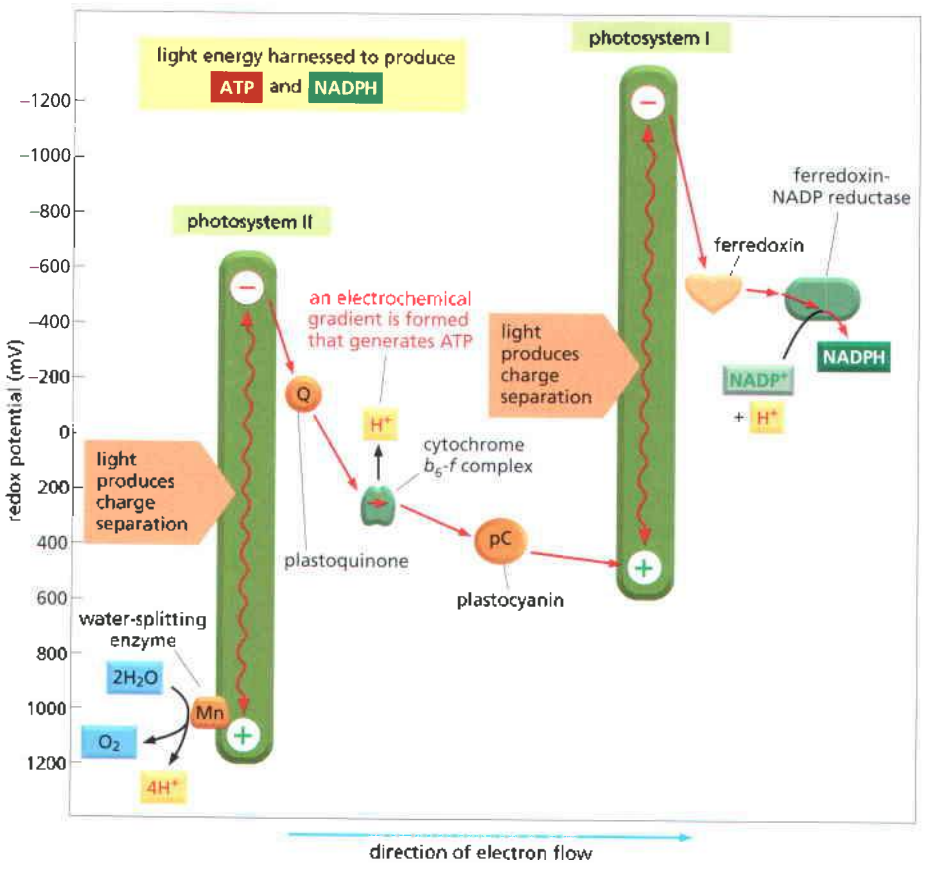


Figure 14–49 Changes in redox potential during photosynthesis. The redox potential for each molecule is indicated by its position along the vertical axis. Note that photosystem II passes electrons derived from water to photosystem I. The net electron flow through the two photosystems in series is from water to $NADP^+$, and it produces NADPH as well as ATP.

The ATP is synthesized by an ATP synthase that harnesses the electrochemical proton gradient produced by the three sites of H^+ activity that are highlighted in Figure 14–48. This Z scheme for ATP production is called noncyclic photophosphorylation, to distinguish it from a cyclic scheme that utilizes only photosystem I (see the text).

A single quantum of visible light does not have enough energy to move an electron all the way from the bottom of photosystem II to the top of photosystem I, which is presumably the energy change required to pass an electron efficiently from water to $NADP^+$. The use of two separate photosystems in series makes the energy from two quanta of light available for this purpose. In addition, there is enough energy left over to enable the electron-transport chain that links the two photosystems to pump H^+ across the thylakoid membrane (or the plasma membrane of cyanobacteria), so that the ATP synthase can harness some of the light-derived energy for ATP production.

Chloroplasts Can Make ATP by Cyclic Photophosphorylation Without Making NADPH

In the noncyclic photophosphorylation scheme just discussed, high-energy electrons leaving photosystem II are harnessed to generate ATP and are passed on to photosystem I to drive the production of NADPH. This produces slightly more than 1 molecule of ATP for every pair of electrons that passes from H_2O to $NADP^+$ to generate a molecule of NADPH. But carbon fixation requires 1.5 molecules of ATP per NADPH molecule (see Figure 14–40). To produce extra ATP, chloroplasts can switch photosystem I into a cyclic mode so that it produces ATP instead of NADPH. In this process, called *cyclic photophosphorylation*, the high-energy electrons from photosystem I are transferred to the cytochrome b_6-f complex rather than being passed on to $NADP^+$. From the b_6-f complex, the electrons are passed back to photosystem I at a low energy. The only net result, besides the conversion of some light energy to heat, is that H^+ is pumped across the thylakoid membrane by the b_6-f complex as electrons pass through it, thereby increasing the electrochemical proton gradient that drives the ATP synthase. (This is analogous to the cyclic process that occurs in purple nonsulfur bacteria in Figure 14–73, below.)

To summarize, cyclic photophosphorylation involves only photosystem I, and it produces ATP without the formation of either NADPH or O_2. The relative activities of cyclic and noncyclic electron flows are regulated by the cell to determine how much light energy is converted into reducing power (NADPH) and how much into high-energy phosphate bonds (ATP).

Photosystems I and II Have Related Structures, and Also Resemble Bacterial Photosystems

The mechanisms of fundamental cell processes such as DNA replication or respiration generally turn out to be the same in eucaryotic cells and in bacteria, even though the number of protein components involved is considerably greater in eucaryotes. Eucaryotes evolved from procaryotes, and the additional proteins were presumably selected for during evolution because they provided an extra degree of efficiency, regulation that was useful to the cell, or both.

Photosystems provide a clear example of this type of evolution. The atomic structures of both eucaryotic photosystems have been revealed by a combination of electron and x-ray crystallography, and the close relationship of photosystem I, photosystem II, and the photochemical reaction center of purple bacteria has been clearly demonstrated from such analyses (**Figure 14–50**).

The Proton-Motive Force Is the Same in Mitochondria and Chloroplasts

The presence of the thylakoid space separates a chloroplast into three rather than the two internal compartments of a mitochondrion. The net effect of H^+ translocation in the two organelles is, however, similar. As illustrated in **Figure 14–51**, in chloroplasts exposed to light, H^+ is pumped out of the stroma (pH 7.5) into the thylakoid space (pH ~4.5), creating a gradient of 3 pH units. This represents a proton-motive force of about 180 mV across the thylakoid membrane,

and it drives ATP synthesis by the ATP synthase embedded in this membrane. The force is nearly the same as that across the inner mitochondrial membrane, but nearly all of it is contributed by the pH gradient rather than by a membrane potential, unlike the case in mitochondria.

For both mitochondria and chloroplasts, the catalytic site of the ATP synthase is at a pH of about 7.5 and is located in a large organelle compartment (matrix or stroma) that is packed full of soluble enzymes. It is here that all of the organelle's ATP is made (see Figure 14–51).

Carrier Proteins in the Chloroplast Inner Membrane Control Metabolite Exchange with the Cytosol

If chloroplasts are isolated in a way that leaves their inner membrane intact, this membrane can be shown to have a selective permeability, reflecting the presence of specific transporters. Most notably, much of the glyceraldehyde 3-phosphate produced by CO_2 fixation in the chloroplast stroma is transported out of the chloroplast by an efficient antiport system that exchanges three-carbon sugar phosphates for an inward flux of inorganic phosphate.

Glyceraldehyde 3-phosphate normally provides the cytosol with an abundant source of carbohydrate, which the cell uses as the starting point for many other biosyntheses—including the production of sucrose for export. But this is not all that this molecule provides. Once the glyceraldehyde 3-phosphate reaches the cytosol, it is readily converted (by part of the glycolytic pathway) to 1,3-phosphoglycerate and then 3-phosphoglycerate (see p. 92), generating one molecule of ATP and one of NADH. (A similar two-step reaction, but working in reverse, forms glyceraldehyde 3-phosphate in the carbon-fixation cycle; see Figure 14–40.) As a result, the export of glyceraldehyde 3-phosphate from the chloroplast provides

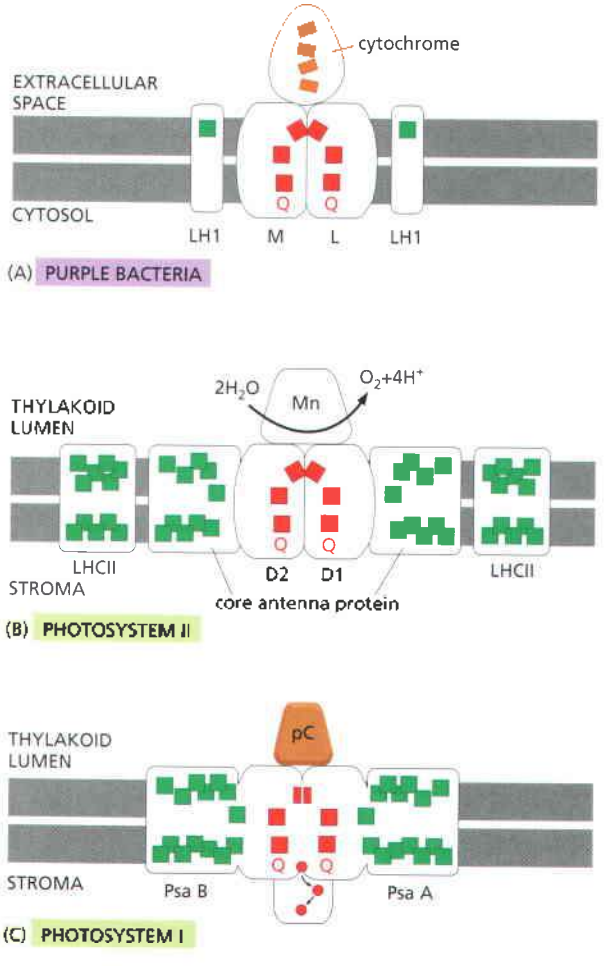


Figure 14–50 Three types of photosynthetic reaction centers compared. Pigments involved in light harvesting are colored *green;* those involved in the central photochemical events are colored *red.* (A) The photochemical reaction center of purple bacteria, whose detailed structure is illustrated in Figure 10–34, contains two related protein subunits, L and M, that bind the pigments involved in the central process illustrated in Figure 14–46. Low-energy electrons are fed into the excited chlorophylls by a cytochrome. LH1 is a protein–pigment complex involved in light harvesting. (B) Photosystem II contains the D_1 and D_2 proteins, which are homologous to the L and M subunits in (A). Low-energy electrons from water are fed into the excited chlorophylls by a manganese cluster. LHCII is a light-harvesting complex that feeds energy into the core antenna proteins (see Figure 14–47). (C) Photosystem I contains the Psa A and Psa B proteins, each of which is equivalent to a fusion of the D_1 or D_2 protein to a core antenna protein of photosystem II. Low-energy electrons are fed into the excited chlorophylls by loosely bound plastocyanin (pC). As indicated, in photosystem I, high-energy electrons are passed from a nonmobile quinone (Q) through a series of three iron-sulfur centers *(red circles).* (Modified from K. Rhee, E. Morris, J. Barber and W. Kühlbrandt, *Nature* 396:283–286, 1998; W. Kühlbrandt, *Nature* 411:896–899, 2001. With permission from Macmillan Publishers Ltd.)

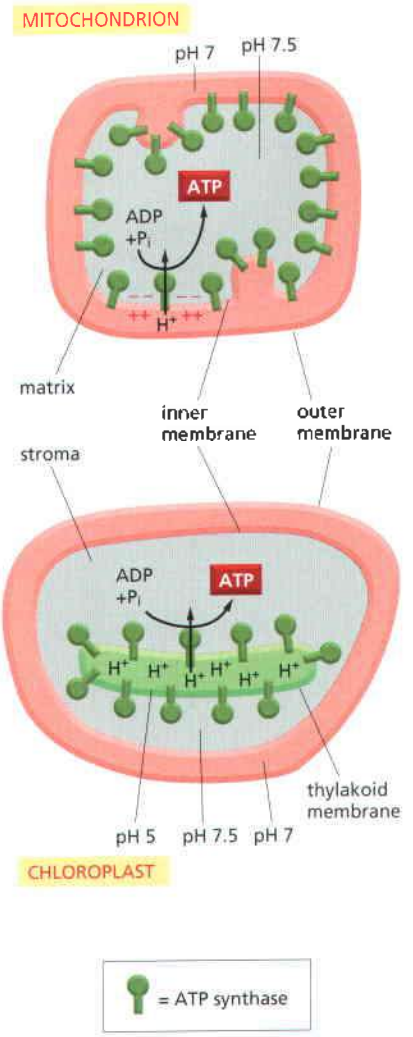


Figure 14–51 A comparison of the flow of H^+ and the orientation of the ATP synthase in mitochondria and chloroplasts. Those compartments with similar pH values have been colored the same. The proton-motive force across the thylakoid membrane consists almost entirely of the pH gradient; a high permeability of this membrane to Mg^{2+} and Cl^- ions allows the flow of these ions to dissipate most of the membrane potential. Mitochondria presumably need a large membrane potential because they could not tolerate having their matrix at pH 10, as would be required to generate their proton-motive force without one.

not only the main source of fixed carbon to the rest of the cell, but also the reducing power and ATP needed for metabolism outside the chloroplast.

Chloroplasts Also Perform Other Crucial Biosyntheses

The chloroplast performs many biosyntheses in addition to photosynthesis. Enzymes in the chloroplast stroma make all of the cell's fatty acids and a set of amino acids, for example. Similarly, the reducing power of light-activated electrons drives the reduction of nitrite (NO_2^-) to ammonia (NH_3) in the chloroplast; this ammonia provides the plant with nitrogen for the synthesis of amino acids and nucleotides. The metabolic importance of the chloroplast for plants and algae therefore extends far beyond its role in photosynthesis.

Summary

Chloroplasts and photosynthetic bacteria obtain high-energy electrons by means of photosystems that capture the electrons that are excited when sunlight is absorbed by chlorophyll molecules. Photosystems are composed of an antenna complex that funnels energy to a photochemical reaction center, where a precisely ordered complex of proteins and pigments allows electron carriers to capture the energy of an excited chlorophyll electron. The most completely understood photochemical reaction center is that of purple photosynthetic bacteria, which contain only a single photosystem. In contrast, there are two distinct photosystems in chloroplasts and cyanobacteria. The two photosystems are normally linked in series, and they transfer electrons from water to $NADP^+$ to form NADPH, with the concomitant production of a transmembrane electrochemical proton gradient. In these linked photosystems, molecular oxygen (O_2) is generated as a by-product of removing four low-energy electrons from two specifically positioned water molecules. The detailed three-dimensional structures of photosystems I and II show a striking degree of homology to the structure of the photosystems of purple photosynthetic bacteria, demonstrating a remarkable degree of conservation over billions of years of evolution.

Compared with mitochondria, chloroplasts have an additional internal membrane (the thylakoid membrane) and a third internal space (the thylakoid space). All electron-transport processes occur in the thylakoid membrane: to make ATP, H^+ is pumped into the thylakoid space, and a backflow of H^+ through an ATP synthase then produces the ATP in the chloroplast stroma. This ATP is used in conjunction with the NADPH made by photosynthesis to drive a large number of biosynthetic reactions in the chloroplast stroma, including the all-important carbon-fixation cycle, which creates carbohydrate from CO_2. Along with some other important chloroplast products, this carbohydrate is exported to the cell cytosol, where—as glyceraldehyde 3-phosphate—it provides organic carbon, ATP, and reducing power to the rest of the cell.

THE GENETIC SYSTEMS OF MITOCHONDRIA AND PLASTIDS

The view that mitochondria and plastids evolved from bacteria that were engulfed by ancestral cells can explain why both types of organelles contain

their own genomes, as well as their own biosynthetic machinery for making RNA and organelle proteins. Mitochondria and plastids are never made from scratch, but instead arise by the growth and division of an existing mitochondrion or plastid. On average, each organelle must double in mass in each cell generation and then be distributed into each daughter cell. Even nondividing cells must replenish organelles that are degraded as part of the continual process of organelle turnover, or produce additional organelles as the need arises. Organelle growth and proliferation is a complicated process because mitochondrial and plastid proteins are encoded in two places: the nuclear genome and the separate genomes harbored in the organelles themselves (Figure 14–52). In Chapter 12, we discuss how selected proteins and lipids are imported into mitochondria and chloroplasts from the cytosol. Here we describe how the organelle genomes are maintained and the contributions they make to organelle biogenesis.

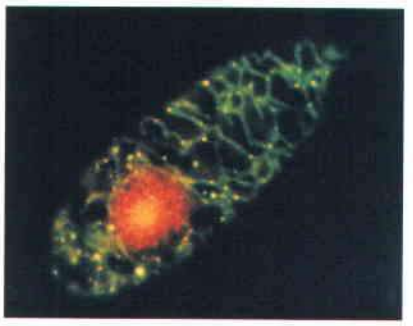


Figure 14–52 Mitochondrial and nuclear DNA stained with a fluorescent dye. This micrograph shows the distribution of the nuclear genome *(red)* and the multiple small mitochondrial genomes (bright *yellow* spots) in a *Euglena gracilis* cell. The DNA is stained with ethidium bromide, a fluorescent dye that emits red light. In addition, the mitochondrial matrix space is stained with a green fluorescent dye that reveals the mitochondria as a branched network extending throughout the cytosol. The superposition of the *green* matrix and the *red* DNA gives the mitochondrial genomes their *yellow* color. (Courtesy of Y. Hayashi and K. Ueda, *J. Cell Sci.* 93:565–570, 1989. With permission from The Company of Biologists.)

Mitochondria and Chloroplasts Contain Complete Genetic Systems

The biosynthesis of mitochondria and plastids requires contributions from two separate genetic systems. Special genes in nuclear DNA encode most of the proteins in mitochondria and chloroplasts. The organelle imports these proteins from the cytosol after they have been synthesized on cytosolic ribosomes. The organelle DNA encodes other organelle proteins that are synthesized on ribosomes within the organelle, using organelle-produced mRNA to specify their amino acid sequence (Figure 14–53). The protein traffic between the cytosol and these organelles seems to be unidirectional, as proteins are normally not exported from mitochondria or chloroplasts to the cytosol. An exception occurs under special conditions when a cell is about to undergo apoptosis. As will be discussed in detail in Chapter 19, the mitochondrion releases intermembrane space proteins (including cytochrome *c*) through its outer mitochondrial membrane as part of an elaborate signaling pathway that is triggered to cause cells to undergo programmed cell death.

The processes of organelle DNA transcription, protein synthesis, and DNA replication (Figure 14–54) take place where the genome is located: in the matrix of mitochondria and the stroma of chloroplasts. Although the proteins that

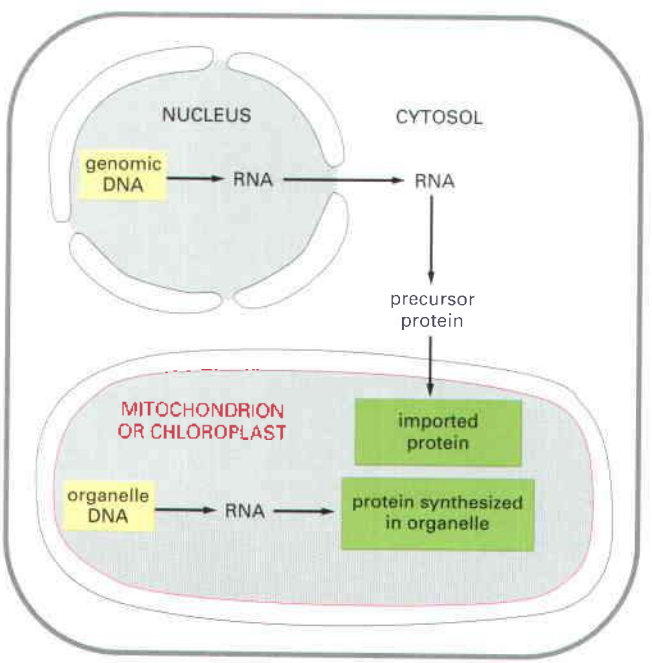


Figure 14–53 The production of mitochondrial and chloroplast proteins by two separate genetic systems. Most of the proteins in these organelles are encoded by the nucleus and must be imported from the cytosol.

mediate these genetic processes are unique to the organelle, the nuclear genome encodes most of them. This is all the more surprising because the protein-synthesis machinery of the organelles resembles that of bacteria rather than that of eucaryotes. The resemblance is particularly close in chloroplasts. For example, chloroplast ribosomes are very similar to *E. coli* ribosomes, both in their structure and in their sensitivity to various antibiotics (such as chloramphenicol, streptomycin, erythromycin, and tetracycline). In addition, protein synthesis in chloroplasts starts with *N*-formyl methionine, as in bacteria, and not with the methionine used for this purpose in the cytosol of eucaryotic cells. Although mitochondrial genetic systems are much less similar to those of present-day bacteria than are the genetic systems of chloroplasts, their ribosomes are also sensitive to antibacterial antibiotics, and protein synthesis in mitochondria also starts with *N*-formyl methionine.

Organelle Growth and Division Determine the Number of Mitochondria and Plastids in a Cell

In mammalian cells, mitochondrial DNA makes up less than 1% of the total cellular DNA. In other cells, however, such as the leaves of higher plants or the very large egg cells of amphibians, a much larger fraction of the cellular DNA may be present in mitochondria or chloroplasts (Table 14–2), and a large fraction of RNA and protein synthesis takes place there.

Mitochondria and plastids are large enough to be observed by light microscopy in living cells. For example, mitochondria can be visualized by expressing a genetically engineered fusion of a mitochondrial protein linked to the green fluorescent protein (GFP) in cells, or cells can be incubated with a fluorescent dye that is specifically taken up by mitochondria because of the electrochemical gradient across their membranes. Such images demonstrate that the mitochondria in living cells are dynamic—frequently dividing, fusing, and changing shape (Figure 14–55), as mentioned previously. Division (fission) and fusion of these organelles are topologically complex processes, because the organelles are enclosed by a double membrane and the mitochondrion must maintain the integrity of its separate mitochondrial compartments during these processes (Figure 14–56).

The number and shape of mitochondria vary dramatically in different cell types and can change in the same cell type under different physiological conditions, ranging from multiple spherical or cylindrically shaped organelles to a single organelle with a branched structure (a reticulum). The arrangement is controlled by the relative rates of mitochondrial division and fusion, which are regulated by dedicated GTPases that reside on mitochondrial membranes. In



Figure 14–54 **An electron micrograph of an animal mitochondrial DNA molecule caught during the process of DNA replication.** The circular DNA genome has replicated only between the two points marked by *red* arrows. The newly synthesized DNA is colored *yellow*. (Courtesy of David A. Clayton.)

Table 14–2 Relative Amounts of Organelle DNA in Some Cells and Tissues

ORGANISM	TISSUE OR CELL TYPE	DNA MOLECULES PER ORGANELLE	ORGANELLES PER CELL	ORGANELLE DNA AS PERCENTAGE OF TOTAL CELLULAR DNA
MITOCHONDRIAL DNA				
Rat	liver	5–10	1000	1
Yeast*	vegetative	2–50	1–50	15
Frog	egg	5–10	10^7	99
CHLOROPLAST DNA				
Chlamydomonas	vegetative	80	1	7
Maize	leaves	0–300**	20–40	0–15**

*The large variation in the number and size of mitochondria per cell in yeasts is due to mitochondrial fusion and fission.
**In maize, the amount of chloroplast DNA drops precipitously in mature leaves, after cell division ceases. the chloroplast DNA is degraded and stable mRNAs persist to provide for protein synthesis.

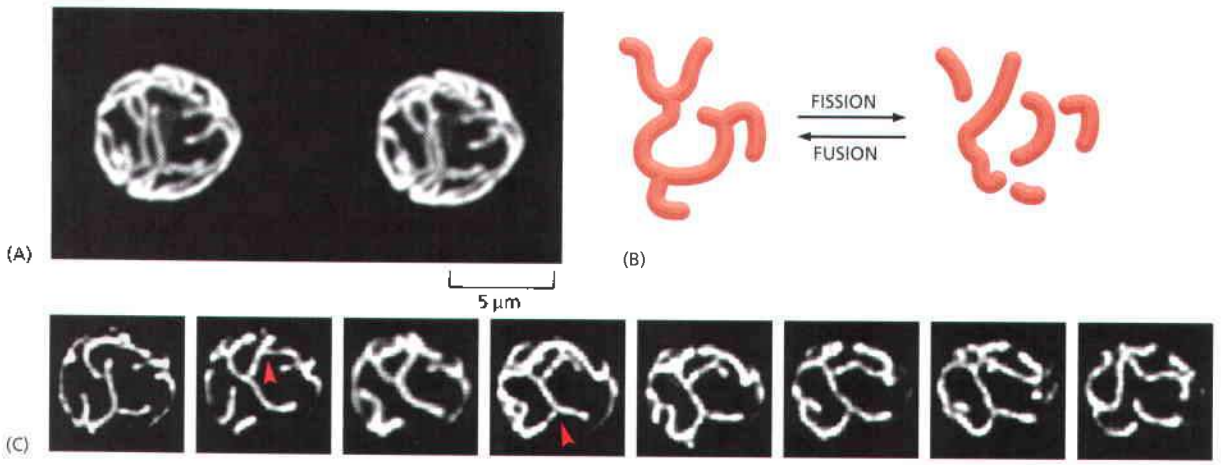
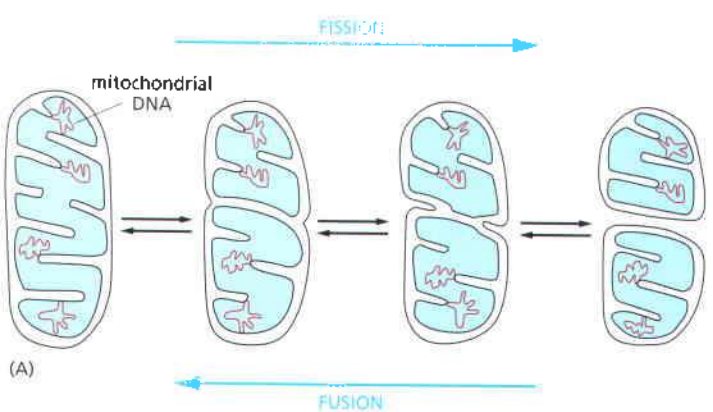


Figure 14–55 A dynamic mitochondrial reticulum. (A) In yeast cells, mitochondria form a continuous reticulum underlying the plasma membrane (stereo pair). (B) A balance between fission and fusion determines the arrangement of the mitochondria in different cells. (C) Time-lapse fluorescent microscopy shows the dynamic behavior of the mitochondrial network in a yeast cell. In addition to shape changes, fission and fusion constantly remodel the network *(red arrows)*. The pictures were taken at 3-minute intervals. (A and C, from J. Nunnari et al., *Mol. Biol. Cell* 8:1233–1242, 1997. With permission by the American Society for Cell Biology.)

addition, the total organelle mass per cell can be regulated according to need. For example, a large increase in mitochondria (as much as 5–10-fold) occurs when a resting skeletal muscle is repeatedly stimulated to contract for a prolonged period.

There can be many copies of the mitochondrial and plastid genomes in the space enclosed by each organelle's inner membrane. The degree of organelle fragmentation determines the number of these genomes present in a single organelle; generally, a single compartment houses many genomes (see Table 14–2). In most cells, the replication of the organelle DNA is not limited to the S phase of the cell cycle, when the nuclear DNA replicates, but occurs throughout the cell cycle—out of phase with cell division. Individual organelle DNA molecules seem to be selected at random for replication, so that in a given cell cycle, some may replicate more than once and others not at all. Nonetheless, under constant conditions, the process is regulated to ensure that the total number of organelle DNA molecules doubles in every cell cycle, as required if each cell type is to maintain a constant amount of organelle DNA.

In special circumstances, the cell can precisely control organelle division. In some algae that contain only one or a few chloroplasts, for example, the organelle divides just before the cell does, in a plane that is identical to the future plane of cell division.



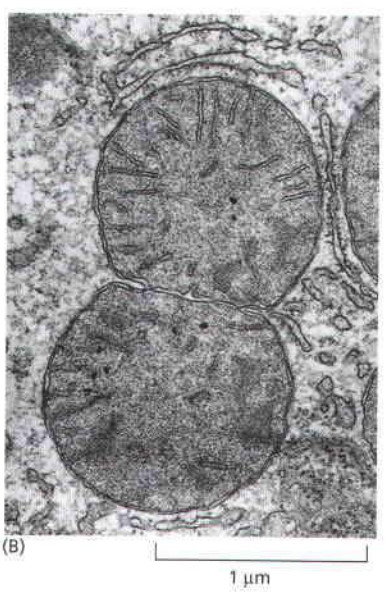


Figure 14–56 Mitochondrial fission and fusion. These processes involve both outer and inner mitochondrial membranes. (A) During fusion and fission, the mitochondrion maintains both its matrix and its intermembrane space compartments. Different membrane fusion machines are thought to operate at the outer and inner membranes. The fission process resembles that of bacterial cell division (discussed in Chapter 16). The pathway shown has been postulated from static views such as that shown in (B). (B) An electron micrograph of a dividing mitochondrion in a liver cell. (B, courtesy of Daniel S. Friend.)

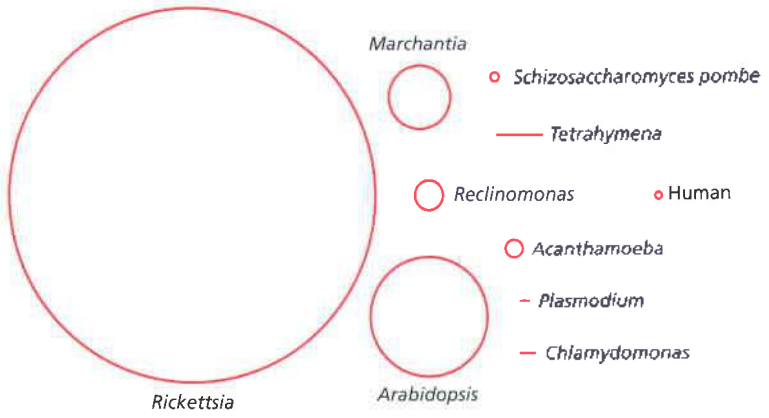


Figure 14–57 Various sizes of mitochondrial genomes. The complete DNA sequences for more than 500 mitochondrial genomes have been determined. The lengths of a few of these mitochondrial DNAs are shown to scale as circles for those genomes thought to be circular and lines for linear genomes. The largest circle represents the genome of *Rickettsia prowazekii*, a small pathogenic bacterium whose genome most closely resembles that of mitochondria. The size of mitochondrial genomes does not correlate well with the number of proteins encoded in them: while human mitochondrial DNA encodes 13 proteins, the 22-fold larger mitochondrial DNA of *Arabidopsis* encodes only 32 proteins—that is, about 2.5-fold as many as human mitochondrial DNA. The extra DNA that is found in *Arabidopsis*, *Marchantia*, and other plant mitochondria may be "junk DNA". The mitochondrial DNA of the protozoan *Reclinomonas americana* has 98 genes, more than the mitochondrion of any other organism analyzed so far. (Adapted from M.W. Gray et al., *Science* 283:1476–1481, 1999. With permission from AAAS.)

Mitochondria and Chloroplasts Have Diverse Genomes

The multiple copies of mitochondrial and chloroplast DNA contained within the matrix or stroma of these organelles are usually distributed in several clusters, called *nucleoids*. Nucleoids are thought to be attached to the inner mitochondrial membrane. The DNA structure in nucleoids is likely to resemble that in bacteria rather than that in eucaryotic chromatin; as in bacteria, for example, there are no histones.

The size range of organelle DNAs is similar to that of viral DNAs. Mitochondrial DNA molecules range in size from less than 6000 nucleotide pairs in *Plasmodium falciparum* (the human malaria parasite) to more than 300,000 nucleotide pairs in some land plants (**Figure 14–57**). The chloroplast genome of land plants ranges in size from 70,000 to 200,000 nucleotide pairs. Chloroplast chromosomes had long been thought to be circular DNA molecules that contain a single copy of the genome, but it is now recognized that such molecules are outnumbered in most chloroplasts by linear DNA molecules formed from multiple copies of the genome, joined end to end. The mixture of structures observed for chloroplast chromosomes is best explained as originating from a DNA replication process that is closely linked to DNA recombination events, analogous to the replication processes observed for viruses such as Herpes. Similar types of mixed structures may also exist for many mitochondrial genomes. However, in mammals, the mitochondrial genome is a simple DNA circle of about 16,500 base pairs (less than 0.001% of the size of the nuclear genome), and it is nearly the same size in animals as diverse as *Drosophila* and sea urchins.

Mitochondria and Chloroplasts Probably Both Evolved from Endosymbiotic Bacteria

The procaryotic character of the organelle genetic systems, especially striking in chloroplasts, suggests that mitochondria and chloroplasts evolved from bacteria that were endocytosed more than 1 billion years ago. According to one version of this *endosymbiont hypothesis*, eucaryotic cells started out as anaerobic organisms without mitochondria or chloroplasts and then established a stable endosymbiotic relation with a bacterium, whose oxidative phosphorylation system they subverted for their own use (**Figure 14–58**). According to this hypothesis, the endocytic event that led to the development of mitochondria occurred when oxygen entered the atmosphere in substantial amounts, more than 1.5×10^9 years ago, before animals and plants separated (see Figure 14–71). The chloroplast seems to have been similarly derived later, from endocytosis of an oxygen-producing photosynthetic organism that resembled a cyanobacterium.

Most of the genes encoding present-day mitochondrial and chloroplast proteins are in the cell nucleus. Thus, an extensive transfer of genes from organelle to nuclear DNA must have occurred during eucaryote evolution. A successful

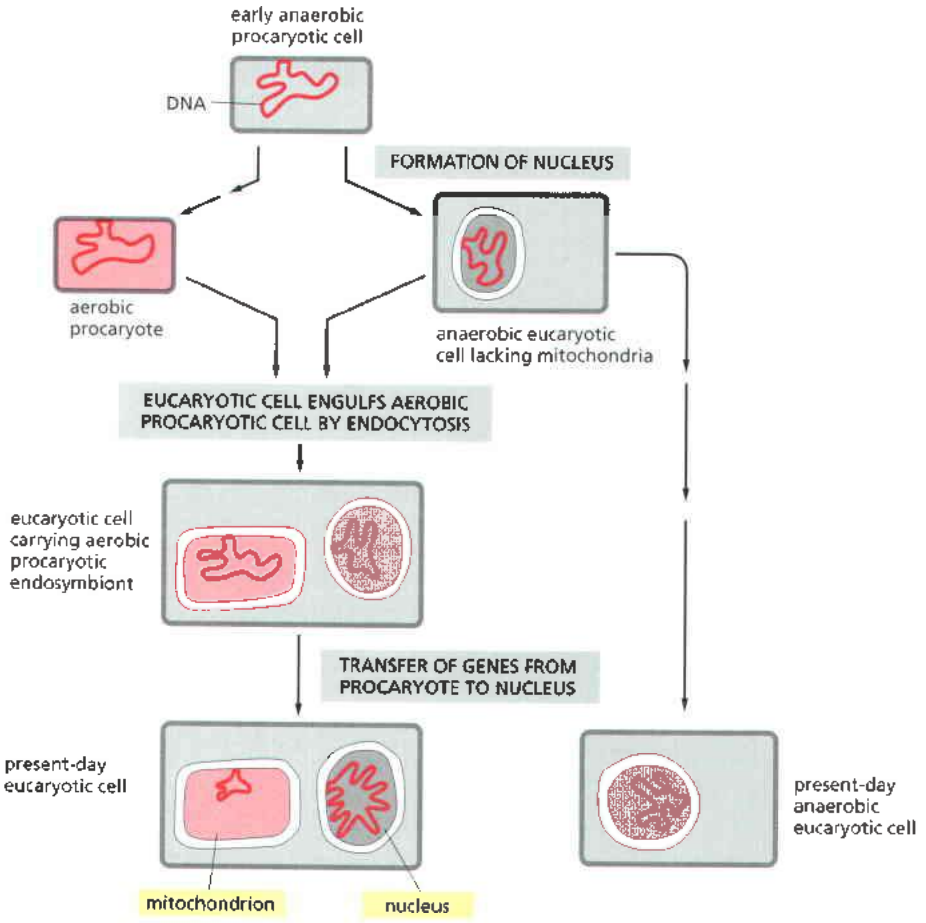


Figure 14–58 One possible evolutionary pathway for the origin of mitochondria. There are anaerobic single-celled eucaryotes present on the Earth today that lack a mitochondrial genome, including the protozoan *Giardia*. However, all those examined appear to contain some type of membrane-enclosed organelle remnant. In these remnants, mitochondrial-like reactions, such as the production of iron-sulfur centers, occur. Thus far, we know of no living example of the anaerobic cell postulated in this figure to have first engulfed the precursors of mitochondria. For this and other reasons, alternative pathways have been postulated in which the combination of a bacterium and an archaea led to both the first eukaryotic cells and the first mitochondria.

transfer of this type is expected to be rare, because a gene moved from organelle DNA needs to change to become a functional nuclear gene: it must adapt to the nuclear and cytoplasmic transcription and translation requirements, and also acquire a signal sequence so that the encoded protein can be delivered to the organelle after its synthesis in the cytosol. Nevertheless, there is evidence that such gene transfers to the nucleus continue to occur in some organisms today.

Gene transfer explains why many of the nuclear genes encoding mitochondrial and chloroplast proteins resemble bacterial genes. The amino acid sequence of the chicken mitochondrial enzyme *superoxide dismutase*, for example, resembles the corresponding bacterial enzyme much more than it resembles the superoxide dismutase found in the cytosol of the same eucaryotic cell.

Gene transfer seems to have been a gradual process. When mitochondrial genomes encoding different numbers of proteins are compared, a pattern of sequential reduction of encoded mitochondrial functions emerges (**Figure 14–59**). The smallest and presumably most highly evolved mitochondrial genomes, for example, encode only a few inner-membrane proteins involved in electron-transport reactions, plus ribosomal RNAs and some tRNAs. Mitochondrial genomes that have remained more complex tend to contain this same subset of genes, plus others. The most complex genomes are characterized by the presence of many extra genes compared with animal and yeast mitochondrial genomes. Many of these genes encode components of the mitochondrial genetic system, such as RNA polymerase subunits and ribosomal proteins; these genes are instead found in the cell nucleus in organisms that have reduced their mitochondrial DNA content.

What type of bacterium gave rise to the mitochondrion? From sequence comparisons, it seems that mitochondria have descended from a particular type of purple photosynthetic bacterium that had previously lost its ability to perform photosynthesis and was left with only a respiratory chain.

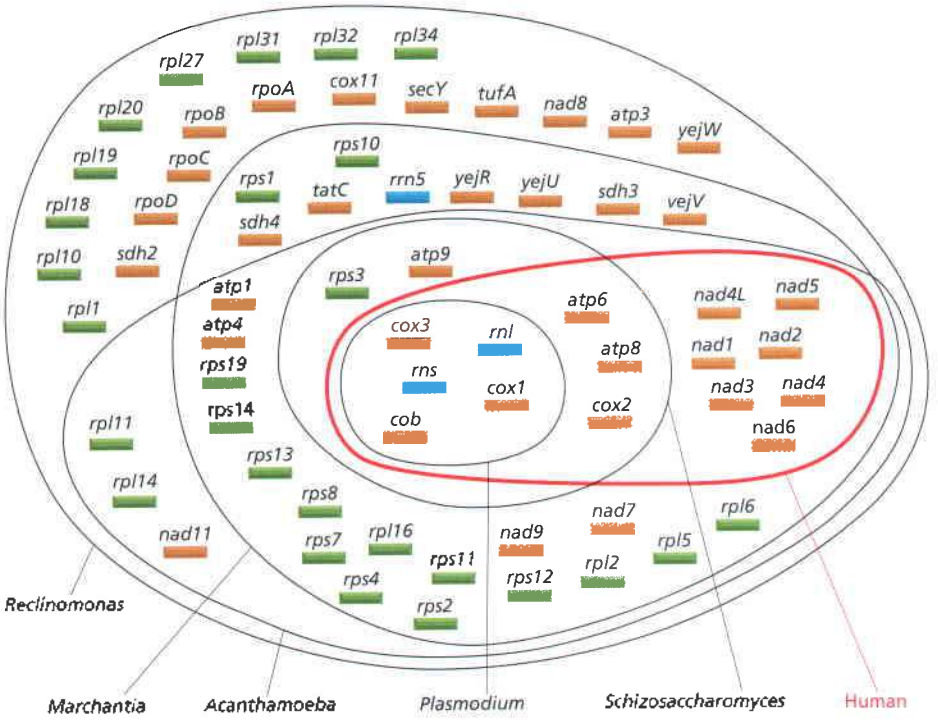


Figure 14–59 Comparison of mitochondrial genomes. Less complex mitochondrial genomes encode subsets of the proteins and ribosomal RNAs that are encoded by larger mitochondrial genomes. There are only four genes present in all known mitochondrial genomes; these encode ribosomal RNAs (*rns* and *rnl*), cytochrome *b* (*cob*), and a cytochrome oxidase subunit (*cox1*). (Adapted from M.W. Gray et al., *Science* 283:1476–1481, 1999. With permission from AAAS.)

Mitochondria Have a Relaxed Codon Usage and Can Have a Variant Genetic Code

The relatively small size of the human mitochondrial genome made it a particularly attractive target for early DNA-sequencing projects, and in 1981, researchers published the complete sequence of its 16,569 nucleotides. By comparing this sequence with known mitochondrial tRNA sequences and with the partial amino acid sequences available for proteins encoded by the mitochondrial DNA, all of the human mitochondrial genes were mapped on the circular DNA molecule (**Figure 14–60**).

Compared with nuclear, chloroplast, and bacterial genomes, the human mitochondrial genome has several surprising features:

1. *Dense gene packing.* Unlike other genomes, nearly every nucleotide seems to be part of a coding sequence, either for a protein or for one of the rRNAs or tRNAs. Since these coding sequences run directly into each other, there is very little room left for regulatory DNA sequences.
2. *Relaxed codon usage.* Whereas 30 or more tRNAs specify amino acids in the cytosol and in chloroplasts, only 22 tRNAs are required for mitochondrial protein synthesis. The normal codon–anticodon pairing rules are relaxed in mitochondria, so that many tRNA molecules recognize any one of the

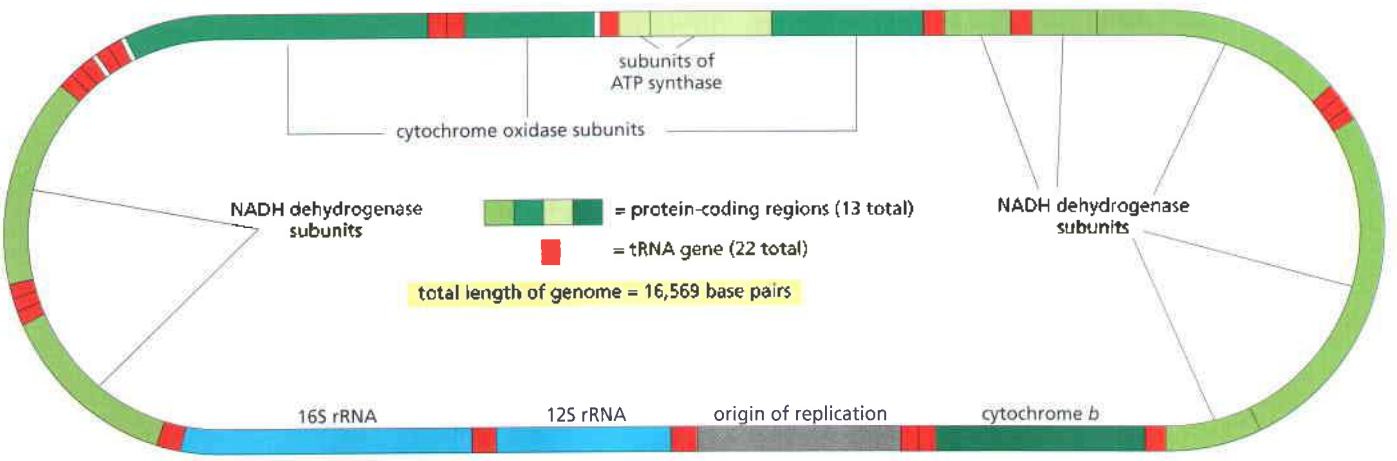


Figure 14–60 The organization of the human mitochondrial genome. The genome contains 2 rRNA genes, 22 tRNA genes, and 13 protein-coding sequences. The DNAs of many other animal mitochondrial genomes have also been completely sequenced. Most of these animal mitochondrial DNAs encode precisely the same genes as humans, with the gene order being identical for animals that range from mammals to fish.

Table 14–3 **Some Differences Between the "Universal" Code and Mitochondrial Genetic Codes***

CODON	"UNIVERSAL" CODE	MITOCHONDRIAL CODES			
		MAMMALS	INVERTEBRATES	YEASTS	PLANTS
UGA	STOP	*Trp*	*Trp*	*Trp*	STOP
AUA	Ile	*Met*	*Met*	*Met*	Ile
CUA	Leu	Leu	Leu	*Thr*	Leu
AGA, AGG	Arg	*STOP*	*Ser*	Arg	Arg

*Red italics indicate that the code differs from the "Universal" code.

four nucleotides in the third (wobble) position. Such "2 out of 3" pairing allows one tRNA to pair with any one of four codons and permits protein synthesis with fewer tRNA molecules.

3. *Variant genetic code.* Perhaps most surprising, comparisons of mitochondrial gene sequences and the amino acid sequences of the corresponding proteins indicate that the genetic code is different: 4 of the 64 codons have different "meanings" from those of the same codons in other genomes (**Table 14–3**).

The close similarity of the genetic code in all organisms provides strong evidence that all cells have evolved from a common ancestor. How, then, do we explain the few differences in the genetic code in many mitochondria? A hint comes from the finding that the mitochondrial genetic code is different in different organisms. In the mitochondrion with the largest number of genes in Figure 14–59, that of the protozoan *Reclinomonas*, the genetic code is unchanged from the standard genetic code of the cell nucleus. Yet UGA, which is a stop codon elsewhere, is read as tryptophan in mitochondria of mammals, fungi, and invertebrates. Similarly, the codon AGG normally codes for arginine, but it codes for *stop* in the mitochondria of mammals and codes for serine in the mitochondria of *Drosophila* (see Table 14–3). Such variation suggests that a random drift can occur in the genetic code in mitochondria. Presumably, the unusually small number of proteins encoded by the mitochondrial genome makes an occasional change in the meaning of a rare codon tolerable, whereas such a change in a large genome would alter the function of many proteins and thereby destroy the cell.

Animal Mitochondria Contain the Simplest Genetic Systems Known

Comparisons of DNA sequences in different organisms reveal that, in vertebrate animals (including ourselves), the rate of nucleotide substitution during evolution has been 10 times greater in the mitochondrial genome than in the nuclear genome. This difference is likely to be due to a reduced fidelity of mitochondrial DNA replication, inefficient DNA repair, or both. Because only about 16,500 DNA nucleotides need to be replicated and expressed as RNAs and proteins in animal cell mitochondria, the error rate per nucleotide copied by DNA replication, maintained by DNA repair, transcribed by RNA polymerase, or translated into protein by mitochondrial ribosomes can be relatively high without damaging one of the relatively few gene products. This could explain why the mechanisms that perform these processes are relatively simple compared with those used for the same purpose elsewhere in cells. The presence of only 22 tRNAs and the unusually small size of the rRNAs (less than two-thirds the size of the *E. coli* rRNAs), for example, would be expected to reduce the fidelity of protein synthesis in mitochondria, although this has not yet been tested adequately.

The relatively high rate of evolution of animal mitochondrial genes makes a comparison of mitochondrial DNA sequences especially useful for estimating the dates of relatively recent evolutionary events, such as the steps in primate evolution.

Some Organelle Genes Contain Introns

The processing of precursor RNAs has an important role in the two mitochondrial systems studied in most detail—human and yeast. In human cells, both strands of the mitochondrial DNA are transcribed at the same rate from a single promoter region on each strand, producing two different giant RNA molecules, each containing a full-length copy of one DNA strand. Transcription is therefore completely symmetric. The transcripts made on one strand are extensively processed by nuclease cleavage to yield the two rRNAs, most of the tRNAs, and about 10 poly-A-containing RNAs. In contrast, the transcript of the other strand is processed to produce only 8 tRNAs and 1 small poly-A-containing RNA; the remaining 90% of this transcript apparently contains no useful information (being complementary to coding sequences synthesized on the other strand) and is degraded. The poly-A-containing RNAs are the mitochondrial mRNAs: although they lack a cap structure at their 5′ end, they carry a poly-A tail at their 3′ end that is added posttranscriptionally by a mitochondrial poly-A polymerase.

Unlike human mitochondrial genes, some plant and fungal (including yeast) mitochondrial genes contain *introns*, which must be removed by RNA splicing. Introns also occur in some in plant chloroplast genes. Many of the introns in organelle genes consist of a family of related nucleotide sequences that are capable of splicing themselves out of the RNA transcripts by RNA-mediated catalysis (discussed in Chapter 6), although proteins generally aid these self-splicing reactions. The presence of introns in organelle genes is surprising, as introns are not common in the genes of the bacteria whose ancestors are thought to have given rise to mitochondria and plant chloroplasts.

In yeasts, the same mitochondrial gene may have an intron in one strain but not in another. Such "optional introns" seem to be able to move in and out of genomes like transposable elements. In contrast, introns in other yeast mitochondrial genes have also been found in a corresponding position in the mitochondria of *Aspergillus* and *Neurospora*, implying that they were inherited from a common ancestor of these three fungi. It is possible that these intron sequences are of ancient origin—tracing back to a bacterial ancestor—and that, although they have been lost from many bacteria, they have been preferentially retained in some organelle genomes where RNA splicing is regulated to help control gene expression.

The Chloroplast Genome of Higher Plants Contains About 120 Genes

More than 20 chloroplast genomes have now been sequenced. The genomes of even distantly related plants (such as tobacco and liverwort) are highly similar, and even those of green algae are closely related (**Figure 14–61**). Chloroplast genes are involved in four main types of processes: transcription, translation, photosynthesis, and the biosynthesis of small molecules such as amino acids, fatty acids, and pigments. Plant chloroplast genes also encode at least 40 proteins whose functions are as yet unknown. Paradoxically, all of the known proteins encoded in the chloroplast are part of larger protein complexes that also contain one or more subunits encoded in the nucleus. We discuss possible reasons for this paradox later.

The genomes of chloroplasts and bacteria have striking similarities. The basic regulatory sequences, such as transcription promoters and terminators, are virtually identical in the two cases. The amino acid sequences of the proteins encoded in chloroplasts are clearly recognizable as bacterial, and several clusters of genes with related functions (such as those encoding ribosomal proteins) are organized in the same way in the genomes of chloroplasts, *E. coli*, and cyanobacteria.

Further comparisons of large numbers of homologous nucleotide sequences should help clarify the exact evolutionary pathway from bacteria to chloroplasts, but some conclusions can already be drawn:

1. Chloroplasts in higher plants arose from photosynthetic bacteria.

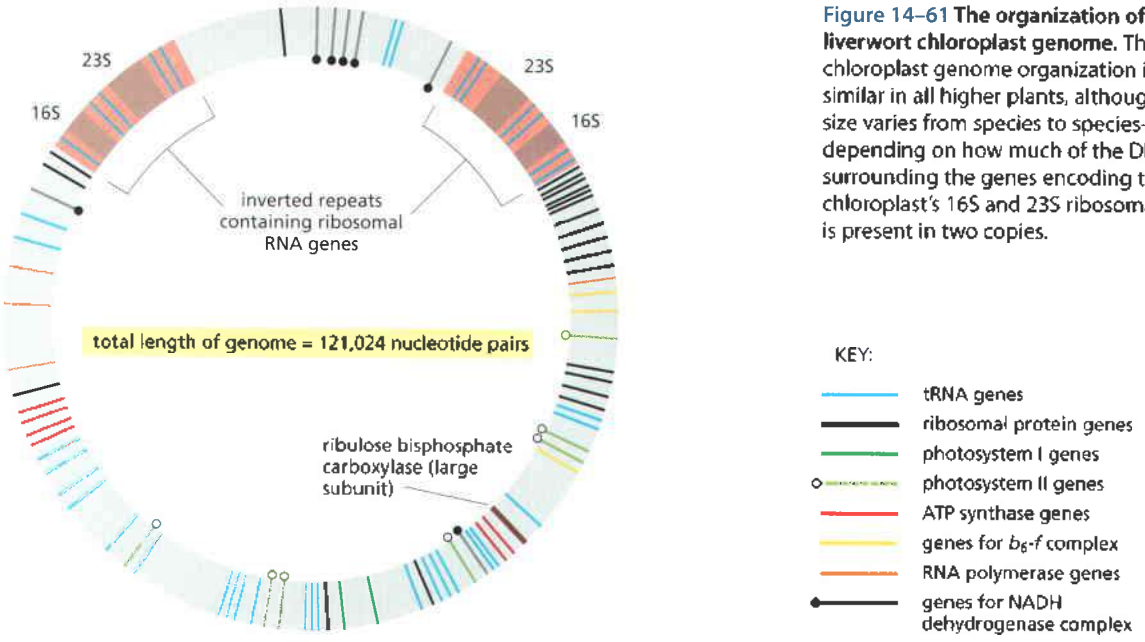


Figure 14–61 The organization of the liverwort chloroplast genome. The chloroplast genome organization is similar in all higher plants, although the size varies from species to species—depending on how much of the DNA surrounding the genes encoding the chloroplast's 16S and 23S ribosomal RNAs is present in two copies.

2. Many of the genes of the original bacterium are now present in the nuclear genome, where they have been integrated and are stably maintained. In higher plants, for example, two-thirds of the 60 or so chloroplast ribosomal proteins are encoded in the cell nucleus; these genes have a clear bacterial ancestry, and the chloroplast ribosomes retain their original bacterial properties.

Mitochondrial Genes Are Inherited by a Non-Mendelian Mechanism

Many experiments on the mechanisms of mitochondrial biogenesis have been performed with *Saccharomyces cerevisiae* (baker's yeast). There are several reasons for this preference. First, when grown on glucose, this yeast has an ability to live by glycolysis alone and can therefore survive with defective mitochondria that cannot perform oxidative phosphorylation. This makes it possible to grow cells with mutations in mitochondrial or nuclear DNA that interfere with mitochondrial function; such mutations are lethal in many other eucaryotes. Second, yeasts are simple unicellular eucaryotes that are easy to grow and characterize biochemically. Finally, these yeast cells normally reproduce asexually by budding, but they can also reproduce sexually. During sexual reproduction two haploid cells mate and fuse to form a diploid zygote, which can either grow mitotically or divide by meiosis to produce new haploid cells.

The ability to control the alternation between asexual and sexual reproduction in the laboratory greatly facilitates genetic analyses. Mutations in mitochondrial genes are not inherited in accordance with the Mendelian rules that govern the inheritance of nuclear genes. Therefore, long before the mitochondrial genome could be sequenced, genetic studies revealed which of the genes involved in yeast mitochondrial function are located in the nucleus and which in the mitochondria. An example of non-Mendelian (cytoplasmic) inheritance of mitochondrial genes in a haploid yeast cell is shown in **Figure 14–62**. In this example, we trace the inheritance of a mutant gene that makes mitochondrial protein synthesis resistant to chloramphenicol.

When a chloramphenicol-resistant haploid cell mates with a chloramphenicol-sensitive wild-type haploid cell, the resulting diploid zygote contains a mixture of mutant and wild-type genomes. The two mitochondrial networks fuse in the zygote, creating one continuous reticulum that contains genomes of both parental cells. When the zygote undergoes mitosis, copies of both mutant and

wild-type mitochondrial DNA are segregated to the diploid daughter cell. In the case of nuclear DNA, each daughter cell receives exactly two copies of each chromosome, one from each parent. By contrast, in the case of mitochondrial DNA, the daughter cell may inherit either more copies of the mutant DNA or more copies of the wild-type DNA. Successive mitotic divisions can further enrich for either DNA, so that subsequently many cells will arise that contain mitochondrial DNA of only one genotype. This stochastic process is called *mitotic segregation.*

When diploid cells that have segregated their mitochondrial genomes in this way undergo meiosis to form four haploid daughter cells, each of the four daughters receives the same mitochondrial genes. This type of inheritance is called non-Mendelian, or *cytoplasmic inheritance,* to contrast it with the Mendelian inheritance of nuclear genes (see Figure 14–62). When non-Mendelian inheritance occurs, it demonstrates that the gene in question is located outside the nuclear chromosomes.

Although clusters of mitochondrial DNA molecules (nucleoids) are relatively immobile in the mitochondrial reticulum because of their anchorage to the inner membrane, individual nucleoids occasionally come together. This is most likely to occur at sites where the two parental mitochondrial networks fuse during zygote formation. When different DNAs are present in the same nucleoid, genetic recombination can occur. This recombination can result in mitochondrial genomes that contain DNA from both parent cells, which are stably inherited after their mitotic segregation.

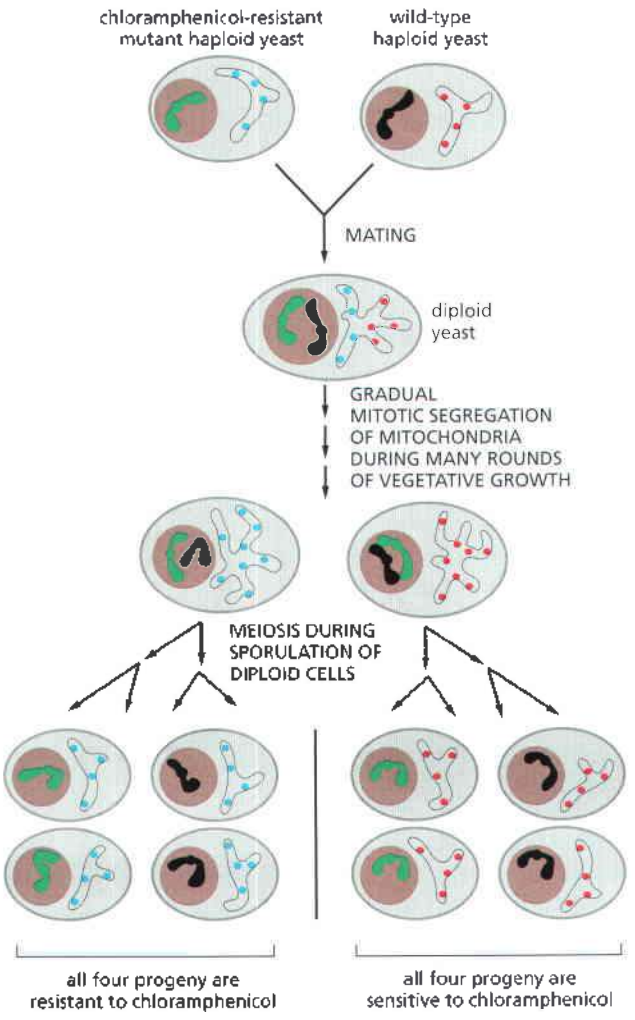


Figure 14–62 The difference in the patterns of inheritance between mitochondrial and nuclear genes of yeast cells. For nuclear genes (Mendelian inheritance), two of the four cells that result from meiosis inherit the gene from one of the original haploid parent cells (*green* chromosomes), and the remaining two cells inherit the gene from the other (*black* chromosomes). By contrast, for mitochondrial genes (non-Mendelian inheritance), it is possible for all four of the cells that result from meiosis to inherit their mitochondrial genes from only one of the two original haploid cells. In this example, the mitochondrial gene is one that, in its mutant form (mitochondrial DNA denoted by *blue dots),* makes protein synthesis in the mitochondrion resistant to chloramphenicol—a protein synthesis inhibitor that acts specifically on the procaryotic-like ribosomes in mitochondria and chloroplasts. We can detect yeast cells that contain the mutant gene by their ability to grow in the presence of chloramphenicol on a substrate, such as glycerol, that cannot be used for glycolysis. With glycolysis blocked, ATP must be provided by functional mitochondria, and therefore the cells that carry the normal (wild-type) mitochondrial DNA *(red dots)* cannot grow.

Organelle Genes Are Maternally Inherited in Many Organisms

The consequences of cytoplasmic inheritance are more profound for some organisms, including ourselves, than they are for yeasts. In yeasts, when two haploid cells mate, they are equal in size and contribute equal amounts of mitochondrial DNA to the zygote (see Figure 14–62). Mitochondrial inheritance in yeasts is therefore *biparental:* both parents contribute equally to the mitochondrial gene pool of the progeny (although, as we have just seen, after several generations of vegetative growth, the *individual* progeny often contain mitochondria from only one parent). In higher animals, by contrast, the egg cell always contributes much more cytoplasm to the zygote than does the sperm. One would therefore expect mitochondrial inheritance in higher animals to be nearly *uniparental*—or, more precisely, maternal. Such *maternal inheritance* has been demonstrated in laboratory animals. A cross between animals carrying type A mitochondrial DNA and animals carrying type B results in progeny that contain only the maternal type of mitochondrial DNA. Similarly, by following the distribution of variant mitochondrial DNA sequences in large families, we find that human mitochondrial DNA is maternally inherited.

In about two-thirds of higher plants, the chloroplasts from the male parent (contained in pollen grains) do not enter the zygote, so that chloroplast as well as mitochondrial DNA is maternally inherited. In other plants, the pollen chloroplasts enter the zygote, making chloroplast inheritance biparental. In such plants, defective chloroplasts are a cause of *variegation:* a mixture of normal and defective chloroplasts in a zygote may sort out by mitotic segregation during plant growth and development, thereby producing alternating green and white patches in leaves. The green patches contain normal chloroplasts, while the white patches contain defective chloroplasts (**Figure 14–63**).

A fertilized human egg carries perhaps 2000 copies of the human mitochondrial genome, all but one or two inherited from the mother. A human in which all of these genomes carried a deleterious mutation would generally not survive. But some mothers carry a mixed population of both mutant and normal mitochondrial genomes. Their daughters and sons inherit this mixture of normal and mutant mitochondrial DNAs and are healthy unless the process of mitotic segregation by chance results in a majority of defective mitochondria in a particular tissue. Muscle and nervous tissues are most at risk, because of their need for particularly large amounts of ATP.

We can identify an inherited disease in humans caused by a mutation in mitochondrial DNA by its passage from affected mothers to both their daughters and their sons, with the daughters but not the sons producing grandchildren with the disease. As expected from the random nature of mitotic segregation, the symptoms of these diseases vary greatly between different family members—including not only the severity and age of onset, but also which tissue is affected.

Consider, for example, the inherited disease *myoclonic epilepsy and ragged red fiber disease (MERRF)*, which can be caused by a mutation in one of the mitochondrial transfer RNA genes. This disease appears when, by chance, a particular tissue inherits a threshold amount of defective mitochondrial DNA genomes. Above this threshold, the pool of defective tRNA decreases the synthesis of the mitochondrial proteins required for electron transport and production of ATP. The result may be muscle weakness or heart problems (from effects on heart muscle), forms of epilepsy or dementia (from effects on nerve cells), or other symptoms. Not surprisingly, a similar variability in phenotypes is found for many other mitochondrial diseases.



Figure 14–63 A variegated leaf. In the white patches, the plant cells have inherited a defective chloroplast. (Courtesy of John Innes Foundation.)

Petite Mutants in Yeasts Demonstrate the Overwhelming Importance of the Cell Nucleus for Mitochondrial Biogenesis

Genetic studies of yeasts have had a crucial role in the analysis of mitochondrial biogenesis. A striking example is provided by studies of yeast mutants that contain large deletions in their mitochondrial DNA, so that all mitochondrial protein synthesis is abolished. Not surprisingly, these mutants cannot make

respiring mitochondria. Some of these mutants lack mitochondrial DNA altogether. Because they form unusually small colonies when grown in medium containing a low concentration of glucose, all mutants with such defective mitochondria are called *cytoplasmic petite mutants.*

Although petite mutants cannot synthesize proteins in their mitochondria and therefore cannot make mitochondria that produce ATP, they nevertheless contain mitochondria. These mitochondria have a normal outer membrane and an inner membrane with poorly developed cristae (**Figure 14–64**). They contain virtually all the mitochondrial proteins that are specified by nuclear genes and imported from the cytosol—including DNA and RNA polymerases, all of the citric acid cycle enzymes, and most inner membrane proteins—demonstrating the overwhelming importance of the nucleus in mitochondrial biogenesis. Petite mutants also show that an organelle that divides by fission can replicate indefinitely in the cytoplasm of proliferating eucaryotic cells, even in the complete absence of its own genome. It is possible that peroxisomes normally replicate in this way (see Figure 12–33).

For chloroplasts, the nearest equivalent to yeast mitochondrial petite mutants are mutants of unicellular algae such as *Euglena.* Mutant algae in which no chloroplast protein synthesis occurs still contain chloroplasts and are perfectly viable if they are provided with oxidizable substrates. If the development of mature chloroplasts is blocked in higher plants, however, either by raising the plants in the dark or because chloroplast DNA is defective or absent, the plants die.

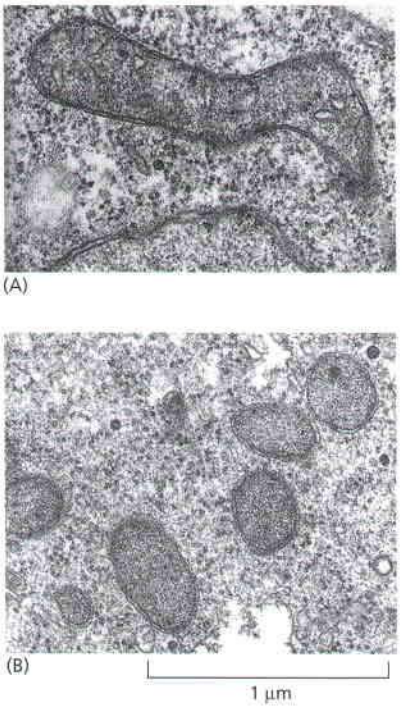


Figure 14–64 Electron micrographs of yeast cells. (A) The structure of normal mitochondria. (B) Mitochondria in a petite mutant. In petite mutants, all the mitochondrion-encoded gene products are missing, and the organelle is constructed entirely from nucleus-encoded proteins. (Courtesy of Barbara Stevens.)

Mitochondria and Plastids Contain Tissue-Specific Proteins That Are Encoded in the Cell Nucleus

Mitochondria can have specialized functions in particular types of cells. The *urea cycle,* for example, is the central metabolic pathway in mammals for disposing of cellular breakdown products that contain nitrogen. These products are excreted in the urine as urea. Nucleus-encoded enzymes in the mitochondrial matrix perform several steps in the cycle. Urea synthesis occurs in only a few tissues, such as the liver, and the required enzymes are synthesized and imported into mitochondria only in these tissues.

The respiratory enzyme complexes in the mitochondrial inner membrane of mammals also contain several tissue-specific, nucleus-encoded subunits that are thought to act as regulators of electron transport. Thus, some humans with a genetic muscle disease have a defective subunit of cytochrome oxidase; since the subunit is specific to skeletal muscle cells, their other cells, including their heart muscle cells, function normally, allowing the individuals to survive. As would be expected, tissue-specific differences are also found among the nucleus-encoded proteins in chloroplasts.

Mitochondria Import Most of Their Lipids; Chloroplasts Make Most of Theirs

The biosynthesis of new mitochondria and chloroplasts requires lipids in addition to nucleic acids and proteins. Chloroplasts tend to make the lipids they require. In spinach leaves, for example, all cellular fatty acid synthesis takes place in the chloroplast, although desaturation of the fatty acids occurs elsewhere. The major glycolipids of the chloroplast are also synthesized locally.

Mitochondria, in contrast, import most of their lipids. Animal cells synthesize the phospholipids phosphatidylcholine and phosphatidylserine in the endoplasmic reticulum and then transfer them to the outer membrane of mitochondria. In addition to decarboxylating imported phosphatidylserine to phosphatidylethanolamine, the main reaction of lipid biosynthesis catalyzed by the mitochondria themselves is the conversion of imported lipids to cardiolipin (bisphosphatidylglycerol). Cardiolipin is a "double" phospholipid that contains four fatty acid tails (**Figure 14–65**). It is found mainly in the mitochondrial inner membrane, where it constitutes about 20% of the total lipid.

We discuss the important question of how specific cytosolic proteins are imported into mitochondria and chloroplasts in Chapter 12.

Mitochondria May Contribute to the Aging of Cells and Organisms

Oxygen gas, O_2, which has a strong affinity for electrons, becomes especially reactive as soon as it acquires a single electron and becomes partially reduced to superoxide (O_2^-). This explains why, at the very end of the respiratory chain, the two oxygen atoms of O_2 remain tightly held by cytochrome oxidase until they have taken up four electrons and can be safely released as two water molecules (see Figure 14–27). As we learn more about the details of biological electron transport processes, we find that evolution has selected for many clever mechanisms that prevent high-energy electrons from escaping from electron transport chains. But accidents nevertheless occur frequently, with a molecule of O_2^- estimated to form in approximately 1 in every 2000 electron transfers from NADH to oxygen.

Cells therefore have backup systems that minimize the damage from superoxide and other reactive oxygen species once they form. Chief among these are *superoxide dismutase (SOD)* enzymes, which catalyze the reaction $O_2^- + O_2^- + 2H^+ \rightarrow H_2O_2 + O_2$. The still toxic hydrogen peroxide product is then further metabolized to water, either by a *catalase* enzyme (which carries out the reaction $H_2O_2 + H_2O_2 \rightarrow 2H_2O + O_2$) or by a *glutathione peroxidase* (which carries out the reaction $H_2O_2 + 2\ GSH \rightarrow 2\ H_2O + GSSG$, where GSH represents the tripeptide glutathione with its cysteine sulfhydryl group).

Approximately 90% of the O_2^- generated in cells is formed inside mitochondria. It is therefore not surprising that mitochondria contain their own superoxide dismutase and glutathione peroxidase enzymes, or that mice deficient in mitochondrial superoxide dismutase (MnSOD) die early in life, unlike mice that lack either the cytosolic or the extracellular versions of this enzyme. Even in normal animals, the mitochondrial DNA is observed to contain a ten-fold higher level of oxidized, abnormal nucleotides than the nuclear DNA. These types of observations, coupled with observations of poorly functioning mitochondria in older human beings, have led to a "vicious cycle" hypothesis that attempts to explain why organisms age. In this view, oxidation damage causes mutations to accumulate in the mitochondrial DNA of somatic tissues, which in turn causes an accelerating rate of oxidation errors in these tissues. This harmful feedback loop continues until the mitochondria of old mice or old humans are producing so many oxidation products that the entire organism decays—giving rise to a debilitating old age.

Is there any truth to this idea? It is clear that aging is a very complex process. Moreover, for lower multicellular organisms such as worms and flies, lifespans can be dramatically elongated through any one of a number of genetic changes. This data suggests that these organisms have been programmed to age and die when they do, rather than wearing out from any accumulation of mutations or damage. As a direct test of the vicious cycle hypothesis, mice engineered to produce half of the normal amount of MnSOD live as long as normal mice and do not appear to age prematurely, even though they accumulate an unusual amount of mitochondrial oxidative damage, show a premature decline in mitochondrial function, and have an increased formation of tumors.

Although the jury is still out with regard to humans, it seems unlikely at present that we shall be able to prolong our lives by simply popping pills of antioxidants.

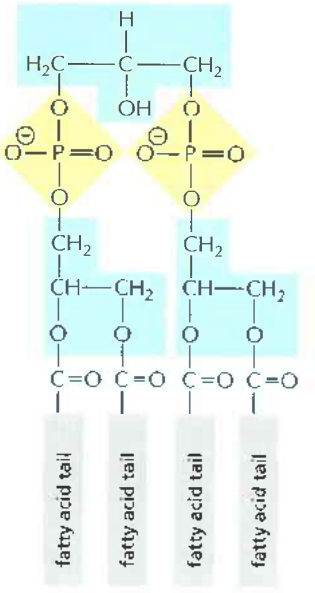


Figure 14–65 **The structure of cardiolipin.** Cardiolipin is an unusual lipid in the inner mitochondrial membrane.

Why Do Mitochondria and Chloroplasts Have Their Own Genetic Systems?

Why do mitochondria and chloroplasts require their own separate genetic systems, when other organelles that share the same cytoplasm, such as peroxisomes and lysosomes, do not? The question is not trivial, because maintaining a

separate genetic system is costly: more than 90 proteins—including many ribosomal proteins, aminoacyl-tRNA synthases, DNA and RNA polymerases, and RNA-processing and RNA-modifying enzymes—must be encoded by nuclear genes specifically for this purpose (**Figure 14–66**). The amino acid sequences of most of these proteins in mitochondria and chloroplasts differ from those of their counterparts in the nucleus and cytosol, and it appears that these organelles have relatively few proteins in common with the rest of the cell. This means that the nucleus must provide at least 90 genes just to maintain each organelle's genetic system.

The reason for such a costly arrangement is not clear. We cannot think of compelling reasons why the proteins made in mitochondria and chloroplasts should be made there rather than in the cytosol.

The individual protein subunits in the various mitochondrial enzyme complexes are highly conserved in evolution, but their site of synthesis is not (see Figure 14–59). Nevertheless, two protein-coding genes, *Cox1* and *Cob*, are present in all mitochondrial genomes. The proteins encoded by these genes are large and hydrophobic, with many membrane-spanning segments, and perhaps they need to be co-translationally inserted into the inner membrane by mitochondrial ribosomes.

Alternatively, the organelle genetic systems may simply be an evolutionary dead-end. In terms of the endosymbiont hypothesis, this would mean that the process whereby the endosymbionts transferred most of their genes to the nucleus stopped before it was complete. Further transfers may have been ruled out, for mitochondria, by alterations in the mitochondrial genetic code that made the remaining mitochondrial genes nonfunctional if they were transferred to the nucleus.

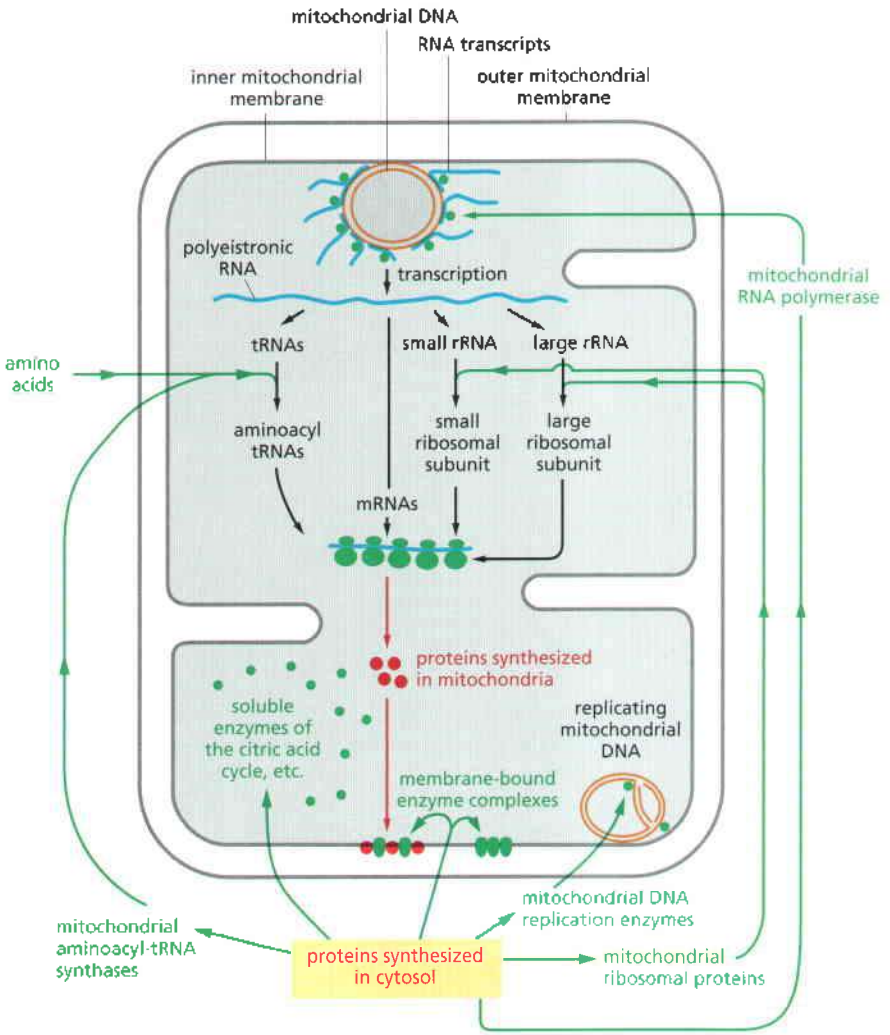


Figure 14–66 The origins of mitochondrial RNAs and proteins. The proteins encoded in the nucleus and imported from the cytosol have a major role in creating the genetic system of the mitochondrion, in addition to contributing most of the organelle's other proteins. Not indicated in this diagram are the additional nucleus-encoded proteins that regulate the expression of individual mitochondrial genes at posttranscriptional levels. The mitochondrion itself contributes only mRNAs, rRNAs, and tRNAs to its genetic system—and in the mitochondria of some organisms, even the tRNAs are imported.

Summary

Mitochondria are additions to cells that allow most eucaryotes to carry out oxidative phosphorylation, while chloroplasts are additions to cell that allow selected eucaryotes (plants and certain algae) to carry out photosynthesis. Presumably as a result of their procaryotic origins, each organelle grows in a coordinated process that requires the contribution of two separate genetic systems—one in the organelle and one in the cell nucleus. Most of the proteins in these organelles are encoded by nuclear DNA, synthesized in the cytosol, and then imported individually into the organelle. Some organelle proteins and RNAs are encoded by the organelle DNA and are synthesized in the organelle itself. The human mitochondrial genome contains about 16,500 nucleotides and encodes 2 ribosomal RNAs, 22 transfer RNAs, and 13 different polypeptide chains. Chloroplast genomes are about 10 times larger and contain about 120 genes. But partly functional organelles form in normal numbers even in mutants that lack a functional organelle genome, demonstrating the overwhelming importance of the nucleus for the biogenesis of both organelles.

The ribosomes of chloroplasts closely resemble bacterial ribosomes, while mitochondrial ribosomes show both similarities and differences that make their origin more difficult to trace. Protein similarities, however, suggest that both organelles originated when a primitive eucaryotic cell entered into a stable endosymbiotic relationship with a bacterium. A purple bacterium is thought to have given rise to the mitochondrion, and (later) a cyanobacterium is thought to have given rise to chloroplasts. Although many of the genes of these ancient bacteria still function to make organelle proteins, most of them have become integrated into the nuclear genome, where they encode bacterialike enzymes that are synthesized on cytosolic ribosomes and then imported into the organelle.

The damage that accumulates in the genomes of mitochondia over time may make a unique contribution to the aging of cells and organisms, although much remains to be done to prove this "vicious cycle" hypothesis.

THE EVOLUTION OF ELECTRON-TRANSPORT CHAINS

Much of the structure, function, and evolution of cells and organisms can be related to their need for energy. We have seen that the fundamental mechanisms for harnessing energy from such disparate sources as light and the oxidation of glucose are the same. Apparently, an effective method for synthesizing ATP arose early in evolution and has since been conserved with only small variations. How did the crucial individual components—ATP synthase, redox-driven H^+ pumps, and photosystems—first arise? Hypotheses about events occurring on an evolutionary time scale are difficult to test. But clues abound, both in the many different primitive electron-transport chains that survive in some present-day bacteria, and in geological evidence about the environment of the Earth billions of years ago.

The Earliest Cells Probably Used Fermentation to Produce ATP

As explained in Chapter 1, the first living cells on Earth are thought to have arisen more than 3×10^9 years ago, when the Earth was not more than about 10^9 years old. The environment lacked oxygen but was presumably rich in geochemically produced organic molecules, and some of the earliest metabolic pathways for producing ATP may have resembled present-day forms of fermentation.

In the process of fermentation, ATP is made by a phosphorylation event that harnesses the energy released when a hydrogen-rich organic molecule, such as glucose, is partly oxidized (see Figure 2–71). The electrons lost from the oxidized organic molecules are transferred (via NADH or NADPH) to a different organic molecule (or to a different part of the same molecule), which thereby becomes

more reduced. At the end of the fermentation process, one or more of the organic molecules produced are excreted into the medium as metabolic waste products; others, such as pyruvate, are retained by the cell for biosynthesis.

The excreted end products are different in different organisms, but they tend to be organic acids (carbon compounds that carry a COOH group). Among the most important of such products in bacterial cells are lactic acid (which also accumulates in the anaerobic glycolysis that occurs in muscle) and formic, acetic, propionic, butyric, and succinic acids.

Electron-Transport Chains Enabled Anaerobic Bacteria to Use Nonfermentable Molecules as Their Major Source of Energy

The early fermentation processes would have provided not only the ATP but also the reducing power (as NADH or NADPH) required for essential biosyntheses. Thus, many of the major metabolic pathways could have evolved while fermentation was the only mode of energy production. With time, however, the metabolic activities of these procaryotic organisms must have changed the local environment, forcing organisms to evolve new biochemical pathways. The accumulation of waste products of fermentation, for example, might have resulted in the following series of changes:

Stage 1. The continuous excretion of organic acids lowered the pH of the environment, favoring the evolution of proteins that function as transmembrane H^+ pumps that can pump H^+ out of the cell to protect it from the dangerous effects of intracellular acidification. One of these pumps may have used the energy available from ATP hydrolysis and could have been the ancestor of the present-day ATP synthase.

Stage 2. At the same time as nonfermentable organic acids were accumulating in the environment and favoring the evolution of an ATP-consuming H^+ pump, the supply of geochemically generated fermentable nutrients, which provided the energy for the pumps and for all other cellular processes, was dwindling. This favored bacteria that could excrete H^+ without hydrolyzing ATP, allowing the ATP to be conserved for other cellular activities. Selective pressures of this kind might have led to the first membrane-bound proteins that could use electron transport between molecules of different redox potentials as the energy source for transporting H^+ across the plasma membrane. Some of these proteins would have found their electron donors and electron acceptors among the nonfermentable organic acids that had accumulated. Present-day bacteria possess many such electron-transport proteins; some bacteria that grow on formic acid, for example, pump H^+ by using the small amount of redox energy derived from the transfer of electrons from formic acid to fumarate (**Figure 14–67**). Others have similar electron-transport components devoted solely to the oxidation and reduction of inorganic substrates (see Figure 14–69, for example).

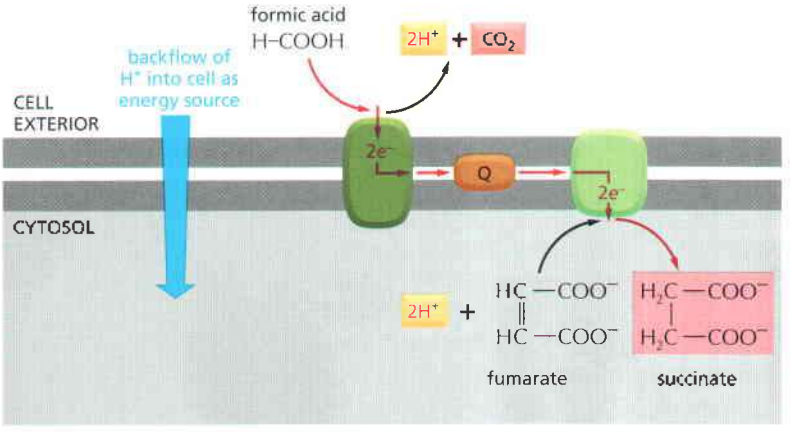
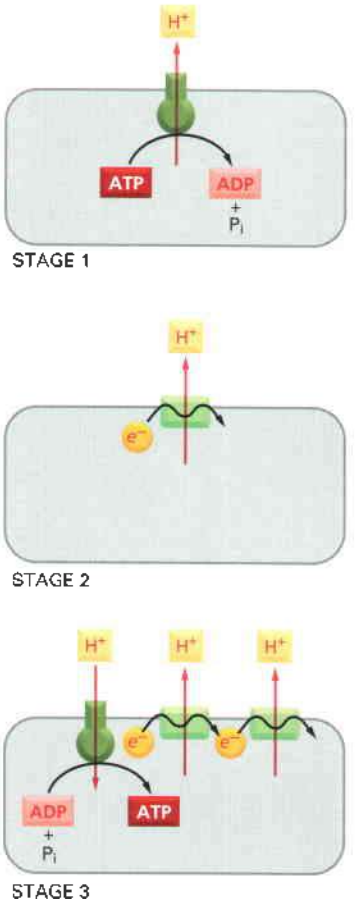


Figure 14–67 The oxidation of formic acid in some present-day bacteria. In such anaerobic bacteria, including *E. coli*, the oxidation is mediated by an energy-conserving electron-transport chain in the plasma membrane. As indicated, the starting materials are formic acid and fumarate, and the products are succinate and CO_2. Note that H^+ is consumed inside the cell and generated outside the cell, which is equivalent to pumping H^+ to the cell exterior. Thus, this membrane-bound electron-transport system can generate an electrochemical proton gradient across the plasma membrane. The redox potential of the formic acid–CO_2 pair is –420 mV, while that of the fumarate–succinate pair is +30 mV. (Compare to Figure 14–29.)

Figure 14–68 The evolution of oxidative phosphorylation mechanisms. One possible sequence is shown; the stages are described in the text.



Stage 3. Eventually some bacteria developed H^+-pumping electron-transport systems that were efficient enough to harness more redox energy than they needed just to maintain their internal pH. Now, bacteria that carried both types of H^+ pumps were at an advantage. In these cells, a large electrochemical proton gradient generated by excessive H^+ pumping allowed protons to leak back into the cell through the ATP-driven H^+ pumps, thereby running them in reverse, so that they functioned as ATP synthases to make ATP. Because such bacteria required much less of the increasingly scarce supply of fermentable nutrients, they proliferated at the expense of their neighbors.

Figure 14–68 shows these three hypothetical stages in the evolution of oxidative phosphorylation mechanisms.

By Providing an Inexhaustible Source of Reducing Power, Photosynthetic Bacteria Overcame a Major Evolutionary Obstacle

The evolutionary steps just described would have solved the problem of maintaining both a neutral intracellular pH and an abundant store of energy, but these steps would not have solved another equally serious problem. The depletion of organic nutrients from the environment meant that organisms had to find some alternative source of carbon to make the sugars that served as the precursors for so many other cellular molecules. Although the CO_2 in the atmosphere provided an abundant potential carbon source, to convert it into an organic molecule such as a carbohydrate requires reducing the fixed CO_2 with a strong electron donor, such as NADH or NADPH, which can provide the high-energy electrons needed to generate each (CH_2O) unit from CO_2 (see Figure 14–40). Early in cellular evolution, strong reducing agents (electron donors) would have been plentiful as products of fermentation. But as the supply of fermentable nutrients dwindled and a membrane-bound ATP synthase began to produce most of the ATP, the plentiful supply of NADH and other reducing agents would have disappeared. It thus became imperative for cells to evolve a new way of generating strong reducing agents.

Presumably, the main reducing agents still available were the organic acids produced by the anaerobic metabolism of carbohydrates, inorganic molecules such as hydrogen sulfide (H_2S) generated geochemically, and water. But the reducing power of these molecules is far too weak to be useful for CO_2 fixation. An early supply of strong electron donors could have been generated by using the electrochemical proton gradient across the plasma membrane to drive a *reverse electron flow*. This would have required the evolution of membrane-bound enzyme complexes resembling an NADH dehydrogenase, and mechanisms of this kind survive in the anaerobic metabolism of some present-day bacteria (**Figure 14–69**).

The major evolutionary breakthrough in energy metabolism, however, was almost certainly the development of photochemical reaction centers that could use the energy of sunlight to produce molecules such as NADH. It is thought that this occurred early in the process of cellular evolution—more than 3×10^9 years ago, in the ancestors of the green sulfur bacteria. Present-day green sulfur bacteria use light energy to transfer hydrogen atoms (as an electron plus a proton) from H_2S to NADPH, thereby creating the strong reducing power required for carbon fixation (**Figure 14–70**). Because the electrons removed from H_2S are at a much more negative redox potential than those of H_2O (– 230 mV for H_2S compared + 820 mV for H_2O), one quantum of light absorbed by the single photosystem in these bacteria is sufficient to achieve a high enough redox potential to generate NADPH via a relatively simple photosynthetic electron-transport chain.

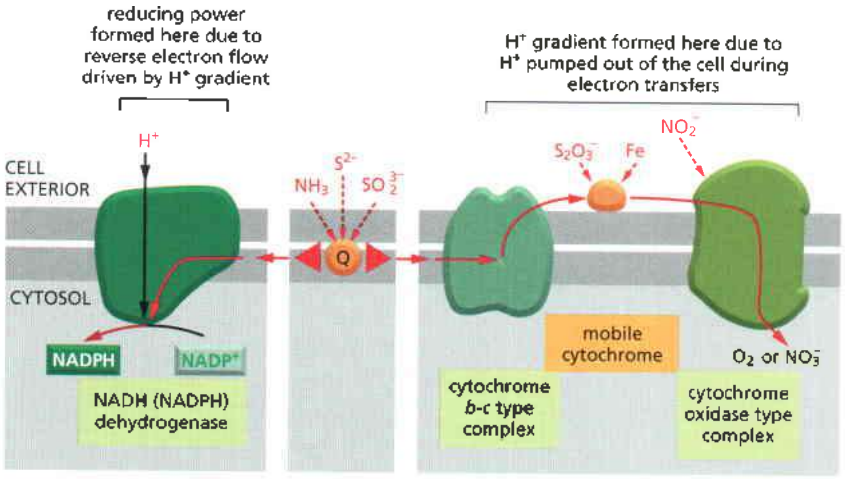


Figure 14–69 Some of the electron-transport pathways in present-day bacteria. Some species can grow anaerobically by substituting nitrate for oxygen as the terminal electron acceptor. Most use the carbon-fixation cycle and synthesize their organic molecules entirely from carbon dioxide. The indicated pathways generate all the ATP and reducing power required from the oxidation of inorganic molecules, such as iron, ammonia, nitrite, and sulfur compounds.

Note that both forward and reverse electron flows occur from the quinone (Q). As in the respiratory chain, the forward electron flows cause H^+ to be pumped out of the cell, and the resulting H^+ gradient drives the production of ATP by an ATP synthase (not shown). The NADPH required for carbon fixation is produced by an energy-requiring reverse electron flow; this energy is also derived from the H^+ gradient, as indicated.

The Photosynthetic Electron-Transport Chains of Cyanobacteria Produced Atmospheric Oxygen and Permitted New Life-Forms

The next step, which is thought to have occurred with the development of the cyanobacteria perhaps 3×10^9 years ago, was the evolution of organisms capable of using water as the electron source for CO_2 reduction. This entailed the evolution of a water-splitting enzyme and also required the addition of a second photosystem, acting in series with the first, to bridge the enormous gap in redox potential between H_2O and NADPH. Present-day structural homologies between photosystems suggest that this change involved the cooperation of a photosystem derived from green bacteria (photosystem I) with a photosystem derived from purple bacteria (photosystem II). The biological consequences of this evolutionary step were far-reaching. For the first time, there were organisms that made only very minimal chemical demands on their environment. These cells could spread and evolve in ways denied to the earlier photosynthetic bacteria, which needed H_2S or organic acids as a source of electrons. Consequently, large amounts of biologically synthesized, reduced organic materials accumulated. Moreover, oxygen entered the atmosphere for the first time.

Oxygen is highly toxic because the oxidation reactions it brings about can randomly alter biological molecules. Many present-day anaerobic bacteria, for example, are rapidly killed when exposed to air. Thus, organisms on the primitive Earth would have had to evolve protective mechanisms against the rising O_2 levels in the environment. Late evolutionary arrivals, such as ourselves, have numerous detoxifying mechanisms that protect our cells from the ill effects of

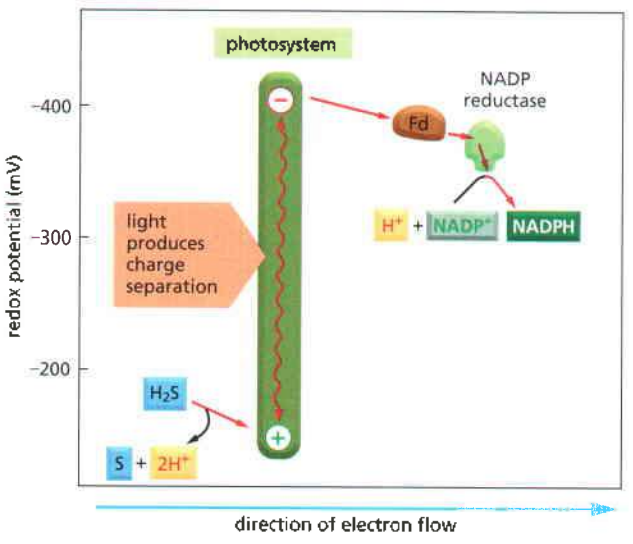


Figure 14–70 The general flow of electrons in a relatively primitive form of photosynthesis observed in present-day green sulfur bacteria. The photosystem in green sulfur bacteria resembles photosystem I in plants and cyanobacteria. Both photosystems use a series of iron–sulfur centers as the electron acceptors that eventually donate their high-energy electrons to ferredoxin (Fd). An example of a bacterium of this type is *Chlorobium tepidum*, which can thrive at high temperatures and low light intensities in hot springs.

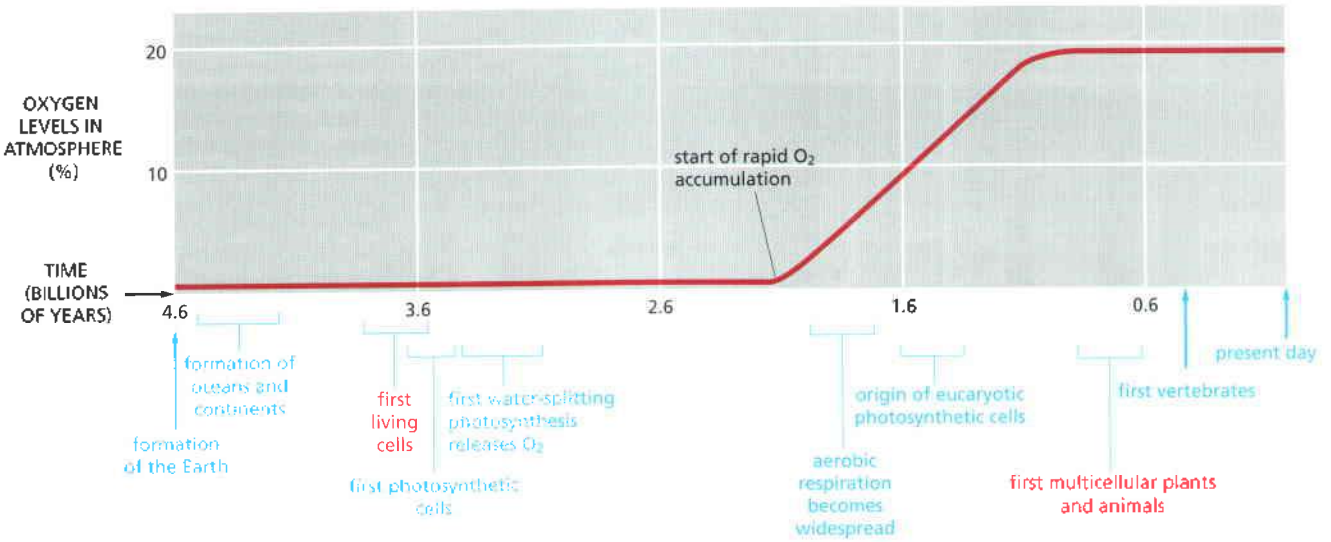


Figure 14–71 Some major events that are believed to have occurred during the evolution of living organisms on Earth. With the evolution of the membrane-based process of photosynthesis, organisms could make their own organic molecules from CO_2 gas. As explained in the text, the delay of more than 10^9 years between the appearance of bacteria that split water and released O_2 during photosynthesis and the accumulation of high levels of O_2 in the atmosphere is thought to be due to the initial reaction of the oxygen with the abundant ferrous iron (Fe^{2+}) that was dissolved in the early oceans. Only when the iron was used up would oxygen have started to accumulate. In response to the rising levels of oxygen in the atmosphere, nonphotosynthetic oxygen-using organisms appeared, and the concentration of oxygen in the atmosphere leveled out.

oxygen. Even so, an accumulation of oxidative damage to our macromolecules has been postulated to be a major cause of human aging, as previously discussed.

The increase in atmospheric O_2 was very slow at first and would have allowed a gradual evolution of protective devices. For example, the early seas contained large amounts of iron in its ferrous oxidation state (Fe^{2+}), and nearly all the O_2 produced by early photosynthetic bacteria could have been utilized in converting Fe^{2+} to Fe^{3+}. This conversion caused the precipitation of huge amounts of ferric oxides, and the extensive banded iron formations in sedimentary rocks beginning about 2.7×10^9 years ago help to date the spread of the cyanobacteria. By about 2×10^9 years ago, the supply of ferrous iron was exhausted, and the deposition of further iron precipitates ceased. Geological evidence suggests that O_2 levels in the atmosphere then began to rise steeply, reaching current levels between 0.5 and 1.5×10^9 years ago (**Figure 14–71**).

The availability of O_2 made possible the development of bacteria that relied on aerobic metabolism to make their ATP. As explained previously, these organisms could harness the large amount of energy released by breaking down carbohydrates and other reduced organic molecules all the way to CO_2 and H_2O. Components of preexisting electron-transport complexes were modified to produce a cytochrome oxidase, so that the electrons obtained from organic or inorganic substrates could be transported to O_2 as the terminal electron acceptor. Depending on the availability of light and O_2, many present-day purple photosynthetic bacteria can switch between photosynthesis and respiration, requiring only relatively minor reorganizations of their electron-transport chains.

As organic materials accumulated on Earth as a result of photosynthesis, some photosynthetic bacteria (including the precursors of *E. coli*) lost their ability to survive on light energy alone and came to rely entirely on respiration. As described previously (see Figure 14–58), it has been suggested that mitochondria first arose more than 1.5×10^9 years ago, perhaps when a primitive eucaryotic cell endocytosed such a respiration-dependent bacterium. Plants are believed to have evolved later, when a descendant of this early aerobic eucaryotic cell endocytosed a photosynthetic bacterium that became the precursor of chloroplasts.

In **Figure 14–72**, we relate these postulated pathways to the various types of bacteria discussed in this chapter. Evolution is always conservative, taking parts of the old and building on them to create something new. Thus, parts of

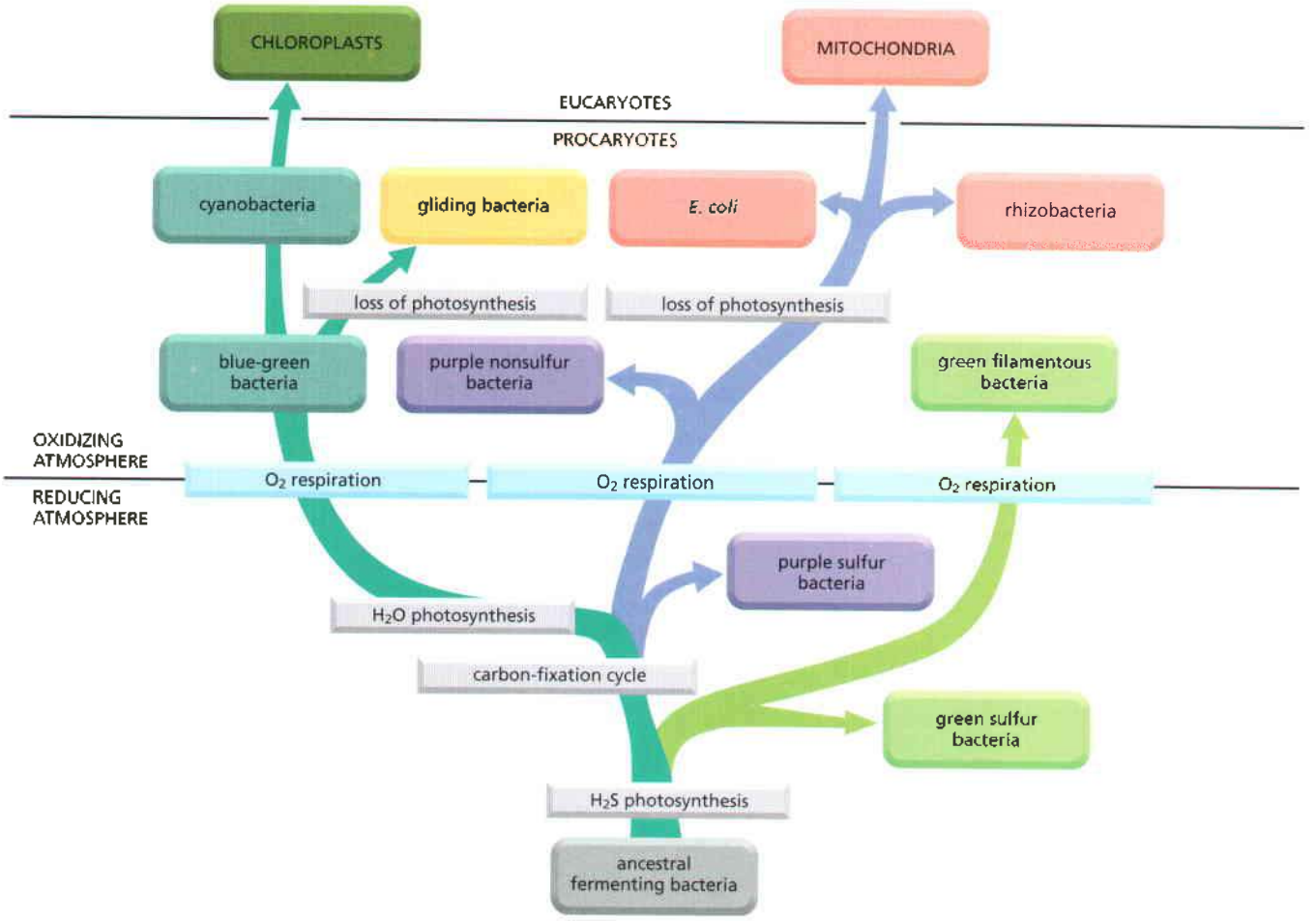


Figure 14–72 A phylogenetic tree of the proposed evolution of mitochondria and chloroplasts and their bacterial ancestors. Oxygen respiration is thought to have begun developing about 2×10^9 years ago. As indicated, it seems to have evolved independently in the green, purple, and blue-green (cyanobacterial) lines of photosynthetic bacteria. It is thought that an aerobic purple bacterium that had lost its ability to photosynthesize gave rise to the mitochondrion, while a cyanobacterium gave rise to chloroplasts. Nucleotide sequence analyses suggest that mitochondria arose from purple bacteria that resembled the rhizobacteria, agrobacteria, and rickettsias—three closely related species known to form intimate associations with present-day eucaryotic cells. Archaea are not known to contain the types of photosystems described in this chapter, and they are not included here.

the electron-transport chains that were derived to service anaerobic bacteria $3–4 \times 10^9$ years ago probably survive, in altered form, in the mitochondria and chloroplasts of today's higher eucaryotes. Consider, for example, the homology in structure and function between the enzyme complex that pumps H^+ in the central segment of the mitochondrial respiratory chain (the cytochrome b-c_1 complex) and its analogs in the electron-transport chains of both bacteria and chloroplasts (**Figure 14–73**).

Summary

Early cells are believed to have been bacteriumlike organisms living in an environment rich in highly reduced organic molecules that had been formed by geochemical processes over the course of hundreds of millions of years. They may have derived most of their ATP by converting these reduced organic molecules to a variety of organic acids, which were then released as waste products. By acidifying the environment, these fermentations may have led to the evolution of the first membrane-bound H^+ pumps, which could maintain a neutral pH in the cell interior by pumping out H^+. The properties of present-day bacteria suggest that an electron-transport-driven H^+ pump and an ATP-driven H^+ pump first arose in this anaerobic environment. Reversal of the ATP-driven pump would have allowed it to function as an ATP synthase. As

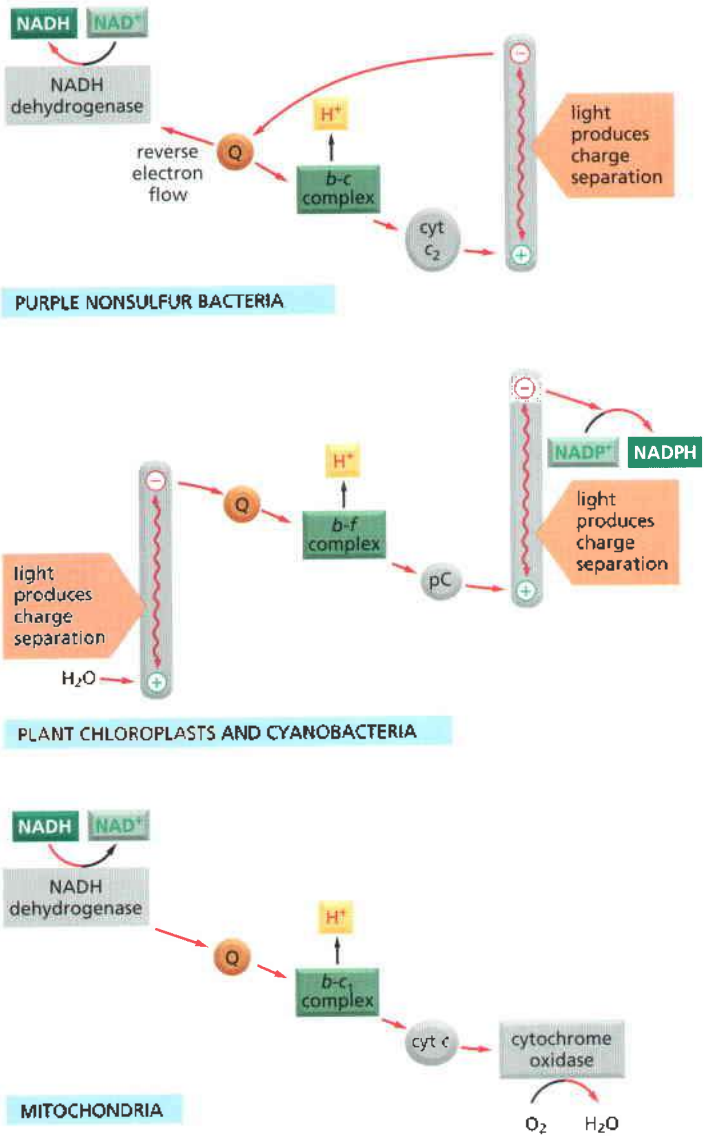


Figure 14–73 A comparison of three electron-transport chains discussed in this chapter. Bacteria, chloroplasts, and mitochondria all contain a membrane-bound enzyme complex that resembles the cytochrome b-c_1 complex of mitochondria. These complexes all accept electrons from a quinone carrier (Q) and pump H^+ across their respective membranes. Moreover, in reconstituted *in vitro* systems, the different complexes can substitute for one another, and the structures of their protein components reveal that they are evolutionarily related.

more effective electron-transport chains developed, the energy released by redox reactions between inorganic molecules and/or accumulated nonfermentable compounds produced a large electrochemical proton gradient, which could be harnessed by the ATP-driven pump for ATP production.

Because preformed organic molecules were replenished only very slowly by geochemical processes, the proliferation of bacteria that used them as the source of both carbon and reducing power could not go on forever. The depletion of fermentable organic nutrients presumably led to the evolution of bacteria that could use CO_2 to make carbohydrates. By combining parts of the electron-transport chains that had developed earlier, light energy was harvested by a single photosystem in photosynthetic bacteria to generate the NADPH required for carbon fixation. The subsequent appearance of the more complex photosynthetic electron-transport chains of the cyanobacteria allowed H_2O to be used as the electron donor for NADPH formation, rather than the much less abundant electron donors required by other photosynthetic bacteria. Life could then proliferate over large areas of the Earth, so that reduced organic molecules accumulated again.

About 2×10^9 years ago, the O_2 released by photosynthesis in cyanobacteria began to accumulate in the atmosphere. Once both organic molecules and O_2 had become abundant, electron-transport chains became adapted for the transport of electrons from NADH to O_2, and efficient aerobic metabolism developed in many bacteria. Exactly the same aerobic mechanisms operate today in the mitochondria of eucaryotes, and there is increasing evidence that both mitochondria and chloroplasts evolved from aerobic bacteria.

PROBLEMS

Which statements are true? Explain why or why not.

14–1 The three respiratory enzyme complexes in the mitochondrial inner membrane exist in structurally ordered arrays that facilitate the correct transfer of electrons between appropriate complexes.

14–2 Lipophilic weak acids short-circuit the normal flow of protons across the inner membrane, thereby eliminating the proton-motive force, stopping ATP synthesis, and blocking the flow of electrons.

14–3 Mutations that are inherited according to Mendelian rules affect nuclear genes; mutations whose inheritance violates Mendelian rules are likely to affect organelle genes.

Discuss the following problems.

14–4 In the 1860s Louis Pasteur noticed that when he added O_2 to a culture of yeast growing anaerobically on glucose, the rate of glucose consumption declined dramatically. Explain the basis for this result, which is known as the Pasteur effect.

14–5 In actively respiring liver mitochondria, the pH inside the matrix is about half a pH unit higher than that in the cytosol. Assuming that the cytosol is at pH 7 and the matrix is a sphere with a diameter of 1 μm [$V = (4/3)\pi r^3$], calculate the total number of protons in the matrix of a respiring liver mitochondrion. If the matrix began at pH 7 (equal to that in the cytosol), how many protons would have to be pumped out to establish a matrix pH of 7.5 (a difference of 0.5 pH units)?

14–6 Heart muscle gets most of the ATP needed to power its continual contractions through oxidative phosphorylation. When oxidizing glucose to CO_2, heart muscle consumes O_2 at a rate of 10 μmol/min per g of tissue, in order to replace the ATP used in contraction and give a steady-state ATP concentration of 5 μmol/g of tissue. At this rate, how many seconds would it take the heart to consume an amount of ATP equal to its steady-state levels? (Complete oxidation of one molecule of glucose to CO_2 yields 30 ATP, 26 of which are derived by oxidative phosphorylation using the 12 pairs of electrons captured in the electron carriers NADH and $FADH_2$.)

14–7 If isolated mitochondria are incubated with a source of electrons such as succinate, but without oxygen, electrons enter the respiratory chain, reducing each of the electron carriers almost completely. When oxygen is then introduced, the carriers become oxidized at different rates (**Figure Q14–1**). How does this result allow you to order the electron carriers in the respiratory chain? What is their order?

14–8 The uncoupler dinitrophenol was once prescribed as a diet drug to aid in weight loss. How would an uncoupler of oxidative phosphorylation promote weight loss? Why do you suppose that it is no longer prescribed?

14–9 How much energy is available in visible light? How much energy does sunlight deliver to the earth? How efficient are plants at converting light energy into chemical energy? The answers to these questions provide an important backdrop to the subject of photosynthesis.

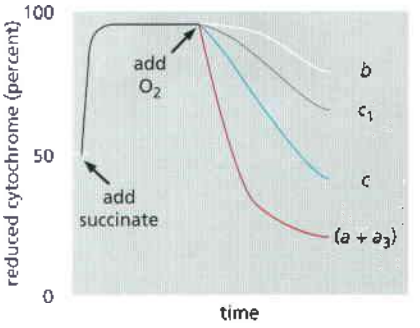


Figure Q14–1 Rapid spectrophotometric analysis of the rates of oxidation of electron carriers in the respiratory chain (Problem 14–7). Cytochromes *a* and a_3 cannot be distinguished and thus are listed as cytochrome $(a + a_3)$.

Each quantum or photon of light has energy $h\nu$, where h is Planck's constant (1.58×10^{-37} kcal sec/photon) and ν is the frequency in sec^{-1}. The frequency of light is equal to c/λ, where c is the speed of light (3.0×10^{17} nm/sec) and λ is the wavelength in nm. Thus, the energy (E) of a photon is

$$E = h\nu = hc/\lambda$$

A. Calculate the energy of a mole of photons (6×10^{23} photons/mole) at 400 nm (violet light), at 680 nm (red light), and at 800 nm (near infrared light).

B. Bright sunlight strikes Earth at the rate of about 0.3 kcal/sec per square meter. Assuming for the sake of calculation that sunlight consists of monochromatic light of wavelength 680 nm, how many seconds would it take for a mole of photons to strike a square meter?

C. Assuming that it takes eight photons to fix one molecule of CO_2 as carbohydrate under optimal conditions (8–10 photons is the currently accepted value), calculate how long it would take a tomato plant with a leaf area of 1 square meter to make a mole of glucose from CO_2. Assume that photons strike the leaf at the rate calculated above and, furthermore, that all the photons are absorbed and used to fix CO_2.

D. If it takes 112 kcal/mole to fix a mole of CO_2 into carbohydrate, what is the efficiency of conversion of light energy into chemical energy after photon capture? Assume again that eight photons of red light (680 nm) are required to fix one molecule of CO_2.

14–10 Recalling Joseph Priestley's famous experiment in which a sprig of mint saved the life of a mouse in a sealed chamber, you decide to do an analogous experiment to see how a C_3 and a C_4 plant do when confined together in a sealed environment. You place a corn plant (C_4) and a geranium (C_3) in a sealed plastic chamber with normal air (300 parts per million CO_2) on a windowsill in your laboratory. What will happen to the two plants? Will they compete or collaborate? If they compete, which one will win and why?

14–11 Examine the variegated leaf shown in **Figure Q14–2**. Yellow patches surrounded by green are common, but there are no green patches surrounded by yellow. Propose an explanation for this phenomenon.



Figure Q14–2 A variegated leaf of *Aucuba japonica* with green and yellow patches (Problem 14–11).

REFERENCES

General

Cramer WA & Knaff DB (1990) Energy Transduction in Biological Membranes: A Textbook of Bioenergetics. New York: Springer-Verlag.

Mathews CK, van Holde KE & Ahern K-G (2000) Biochemistry, 3rd ed. San Francisco: Benjamin Cummings.

Nicholls DG & Ferguson SJ (2002) Bioenergetics, 3rd ed. London: Academic Press.

The Mitochondrion

Abrahams JP, Leslie AG, Lutter R & Walker JE (1994) Structure at 2.8Å resolution of F1-ATPase from bovine heart mitochondria. *Nature* 370:621–628.

Berg HC (2003) The rotary motor of bacterial flagella. *Annu Rev Biochem* 72:19–54.

Boyer PD (1997) The ATP synthase—a splendid molecular machine. *Annu Rev Biochem* 66:717–749.

Ernster L & Schatz G (1981) Mitochondria: a historical review. *J Cell Biol* 91:227s–255s.

Frey TG, Renken CW & Perkins GA (2002) Insight into mitochondrial structure and function from electron tomography. *Biochim Biophys Acta* 1555:196–203.

Pebay-Peyroula E & Brandolin G (2004) Nucleotide exchange in mitochondria: insight at a molecular level. *Curr Opin Struct Biol* 14:420–425.

Meier T, Polzer P, Diederichs K et al (2005) Structure of the rotor ring of F-type Na^+-ATPase from *Ilyobacter tartaricus*. *Science* 308:659–662.

Mitchell P (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* 191:144–148.

Nicholls DG (2006) The physiological regulation of uncoupling proteins. *Biochim Biophys Acta* 1757:459–466.

Racker E & Stoeckenius W (1974) Reconstitution of purple membrane vesicles catalyzing light-driven proton uptake and adenosine triphosphate formation. *J Biol Chem* 249:662–663.

Saraste M (1999) Oxidative phosphorylation at the fin de siecle. *Science* 283:1488–1493.

Scheffler IE (1999) Mitochondria. New York/Chichester: Wiley-Liss.

Stock D, Gibbons C, Arechaga I et al. (2000) The rotary mechanism of ATP synthase. *Curr Opin Struct Biol* 10:672–679.

Weber J (2007) ATP synthase—the structure of the stator stalk. *Trends Biochem Sci* 32:53–56.

Electron-Transport Chains and Their Proton Pumps

Beinert H, Holm RH & Munck E (1997) Iron-sulfur clusters: nature's modular, multipurpose structures. *Science* 277:653–659.

Berry EA, Guergova-Kuras M, Huang LS & Crofts AR (2000) Structure and function of cytochrome *bc* complexes. *Annu Rev Biochem* 69:1005–1075.

Brand MD (2005) The efficiency and plasticity of mitochondrial energy transduction. *Biochem Soc Trans* 33:897–904.

Brandt U (2006) Energy converting NADH:quinone oxidoreductase (complex I). *Annu Rev Biochem* 75:69–92.

Chance B & Williams GR (1955) A method for the localization of sites for oxidative phosphorylation. *Nature* 176:250–254.

Gottschalk G (1997) Bacterial Metabolism, 2nd ed. New York: Springer.

Gray HB & Winkler JR (1996) Electron transfer in proteins. *Annu Rev Biochem* 65:537–556.

Keilin D (1966) The History of Cell Respiration and Cytochromes. Cambridge: Cambridge University Press.

Sazanov LA (2007) Respiratory complex I: mechanistic and structural insights provided by the crystal structure of the hydrophilic domain. *Biochemistry* 46:2275–2288.

Subramaniam S & Henderson R (2000) Molecular mechanism of vectorial proton translocation by bacteriorhodopsin. *Nature* 406:653–657.

Tsukihara T, Aoyama H, Yamashita E et al (1996) The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8Å. *Science* 272:1136–1144.

Wikstrom M & Verkhovsky MI (2006) Towards the mechanism of proton pumping by the haem-copper oxidases. *Biochim Biophys Acta* 1757:1047–1051.

Chloroplasts and Photosynthesis

Blankenship RE (2002) Molecular Mechanisms of Photosynthesis. Oxford, UK:Blackwell Scientific.

Bassham JA (1962) The path of carbon in photosynthesis. *Sci Am* 206:88–100.

Deisenhofer J & Michel H (1989) Nobel lecture. The photosynthetic reaction centre from the purple bacterium *Rhodopseudomonas viridis*. *EMBO J* 8:2149–2170.

Edwards GE, Furbank RT, Hatch MD & Osmond CB (2001) What does it take to be c(4)?—lessons from the evolution of c(4) photosynthesis. *Plant Physiol* 125:46–49.

Ferreira KN, Iverson TM, Maghlaoui K et al (2004) Architecture of the photosynthetic oxygen-evolving center. *Science* 303:1831–1838.

Iwata S & Barber J (2004) Structure of photosystem II and molecular architecture of the oxygen-evolving centre. *Curr Opin Struct Biol* 14:447–453.

Jordan P, Fromme P, Witt HT et al (2001) Three-dimensional structure of cyanobacterial photosystem I at 2.5Å resolution. *Nature* 411:909–917.

Merchant S & Sawaya MR (2005) The light reactions: A guide to recent acquisitions for the picture gallery. *The Plant Cell* 17:648–663.

Nelson N & Ben-Shem A (2004) The complex architecture of oxygenic photosynthesis. *Nature Rev Mol Cell Biol* 5:971–982.

The Genetic Systems of Mitochondria and Plastids

Anderson S, Bankier AT, Barrell BG et al (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465.

Balaban RS, Nemoto S & Finkel T (2005) Mitochondria, oxidants, and aging. *Cell* 120:483–495.

Bendich AJ (2004) Circular chloroplast genomes:The grand illusion. *The Plant Cell* 16:1661–1666.

Birky CW, Jr (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc Natl Acad Sci USA* 92:11331–11338.

Bullerwell CE & Gray MW (2004) Evolution of the mitochondrial genome: protist connections to animals, fungi and plants. *Curr Opin Microbiol* 7:528–534.

Cavalier-Smith T (2002) Chloroplast evolution: secondary symbiogenesis and multiple losses. *Curr Biol* 12:R62–R64.

Chen XJ & Butow RA (2005) The organization and inheritance of the mitochondrial genome. *Nature Rev Genet* 6:815–825.

Clayton DA (2000) Vertebrate mitochondrial DNA—a circle of surprises. *Exp Cell Res* 255:4–9.

Daly DO & Whelan J (2005) Why genes persist in organelle genomes. *Genome Biol* 6:110.

de Duve C (2007) The origin of eukaryotes: A reappraisal. *Nature Rev Genet* 8:395–403.

Dyall SD, Brown MT & Johnson PJ (2004) Ancient invasions: From endosymbionts to organelles. *Science* 304:253–257.

Hoppins S, Lackner L & Nunnari J (2007) The Machines that Divide and Fuse Mitochondria. *Annu Rev Biochem* 76:751–80.

Wallace DC (1999) Mitochondrial diseases in man and mouse. *Science* 283:1482–1488.

The Evolution of Electron-Transport Chains

Blankenship RE & Bauer CE (eds) (1995) Anoxygenic Photosynthetic Bacteria. Dordrecht: Kluwer.

de Las Rivas J, Balsera M & Barber J (2004) Evolution of oxygenic photosynthesis: genome-wide analysis of the OEC extrinsic proteins. *Trends Plant Sci* 9:18–25.

Orgel LE (1998) The origin of life—a review of facts and speculations. *Trends Biochem Sci* 23:491–495.

Schafer G, Purschke W & Schmidt CL (1996) On the origin of respiration: electron transport proteins from archaea to man. *FEMS Microbiol Rev* 18:173–188.

Skulachev VP (1994) Bioenergetics: the evolution of molecular mechanisms and the development of bioenergetic concepts. *Antonie Van Leeuwenhoek* 65:271–284.

Mechanisms of Cell Communication

15

To make a multicellular organism, cells must communicate, just as humans must communicate if they are to organize themselves into a complex society. And just as human communication involves more than the passage of noises from mouth to ear, so cell–cell communication involves more than the transmission of chemical signals across the space between one cell and another. Complex intracellular mechanisms are needed to control which signals are emitted at what time and to enable the signal-receiving cell to interpret those signals and use them to guide its behavior. According to the fossil record, sophisticated multicellular organisms did not appear on Earth until unicellular organisms resembling present-day procaryotes had already been in existence for about 2.5 billion years. The long delay may reflect the difficulty of evolving the language systems of animal, plant, and fungal cells—the machinery that would enable cells sharing the same genome to collaborate and coordinate their behavior, specializing in different ways and subordinating their individual chances of survival to the interests of the multicellular organism as a whole. These highly evolved mechanisms of cell–cell communication are the topic of this chapter.

Communication between cells is mediated mainly by **extracellular signal molecules**. Some of these operate over long distances, signaling to cells far away; others signal only to immediate neighbors. Most cells in multicellular organisms both emit and receive signals. Reception of the signals depends on *receptor proteins*, usually (but not always) at the cell surface, which bind the signal molecule. The binding activates the receptor, which in turn activates one or more *intracellular signaling pathways*. These relay chains of molecules—mainly *intracellular signaling proteins*—process the signal inside the receiving cell and distribute it to the appropriate intracellular targets. These targets are generally *effector proteins*, which are altered when the signaling pathway is activated and implement the appropriate change of cell behavior. Depending on the signal and the nature and state of the receiving cell, these effectors can be gene regulatory proteins, ion channels, components of a metabolic pathway, or parts of the cytoskeleton—among other things (**Figure 15–1**).

We begin this chapter by discussing the general principles of cell communication. We then consider, in turn, the main families of cell-surface receptor proteins and the principal intracellular signaling pathways they activate. The main focus of the chapter is on animal cells, but we end by considering the special features of cell communication in plants.

In This Chapter

GENERAL PRINCIPLES OF CELL COMMUNICATION 879

SIGNALING THROUGH G-PROTEIN-COUPLED CELL-SURFACE RECEPTORS (GPCRs) AND SMALL INTRACELLULAR MEDIATORS 904

SIGNALING THROUGH ENZYME-COUPLED CELL-SURFACE RECEPTORS 921

SIGNALING PATHWAYS DEPENDENT ON REGULATED PROTEOLYSIS OF LATENT GENE REGULATORY PROTEINS 946

SIGNALING IN PLANTS 955

GENERAL PRINCIPLES OF CELL COMMUNICATION

Long before multicellular organisms appeared on Earth, unicellular organisms had developed mechanisms for responding to physical and chemical changes in their environment. These almost certainly included mechanisms for response to the presence of other cells. Evidence comes from studies of present-day unicellular organisms such as bacteria and yeasts. Although these cells largely lead independent lives, they can communicate and influence one another's behavior. Many bacteria, for example, respond to chemical signals that are secreted by their neighbors and increase in concentration with increasing population density. This

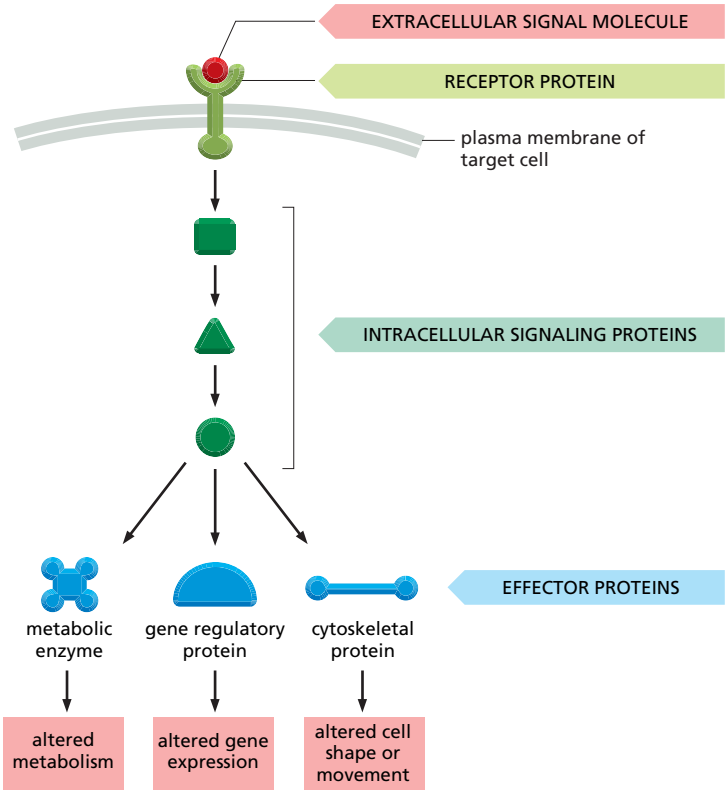


Figure 15–1 A simple intracellular signaling pathway activated by an extracellular signal molecule. The signal molecule usually binds to a receptor protein that is embedded in the plasma membrane of the target cell and activates one or more intracellular signaling pathways mediated by a series of signaling proteins. Finally, one or more of the intracellular signaling proteins alters the activity of effector proteins and thereby the behavior of the cell.

process, called *quorum sensing*, allows bacteria to coordinate their behavior, including their motility, antibiotic production, spore formation, and sexual conjugation.

Similarly, yeast cells communicate with one another in preparation for mating. The budding yeast *Saccharomyces cerevisiae* provides a well-studied example: when a haploid individual is ready to mate, it secretes a peptide *mating factor* that signals cells of the opposite mating type to stop proliferating and prepare to mate (**Figure 15–2**). The subsequent fusion of two haploid cells of opposite mating type produces a diploid cell, which can then undergo meiosis and sporulate, generating haploid cells with new assortments of genes (see Figure 21–3B). The reshuffling of genes through sexual reproduction helps a species survive in an unpredictably variable environment (as discussed in Chapter 21).

Studies of yeast mutants that are unable to mate have identified many proteins that are required in the signaling process. These proteins form a signaling network that includes cell-surface receptor proteins, GTP-binding proteins, and protein kinases, and each of these categories has close relatives among the receptors and intracellular signaling proteins in animal cells. Through gene duplication and divergence, however, the signaling systems in animals have become much more elaborate than those in yeasts; the human genome, for example, contains more than 1500 genes that encode receptor proteins, and the number of different receptor proteins is further increased by alternative RNA splicing and post-translational modifications.

The large numbers of signal proteins, receptors, and intracellular signaling proteins used by animals can be grouped into a much smaller number of protein families, most of which have been highly conserved in evolution. Flies, worms, and mammals all use essentially similar machinery for cell communication, and many of the key components and signaling pathways were first discovered through analysis of mutations in *Drosophila* and *C. elegans*.

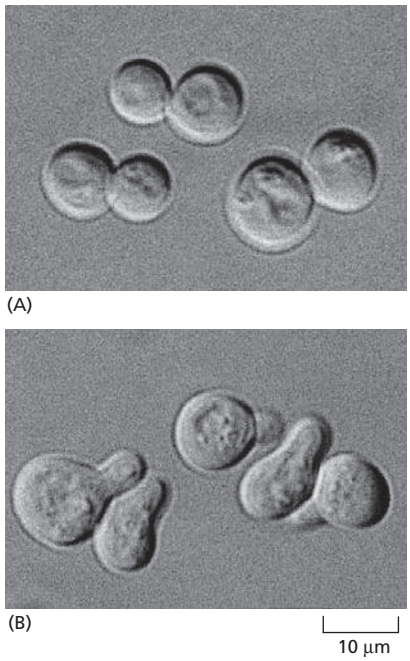


Figure 15–2 Budding yeast cells responding to a mating factor. (A) The cells are normally spherical. (B) In response to the mating factor secreted by neighboring yeast cells, they put out a protrusion toward the source of the factor in preparation for mating. (Courtesy of Michael Snyder.)

Extracellular Signal Molecules Bind to Specific Receptors

Cells in multicellular animals communicate by means of hundreds of kinds of signal molecules. These include proteins, small peptides, amino acids,

nucleotides, steroids, retinoids, fatty acid derivatives, and even dissolved gases such as nitric oxide and carbon monoxide. Most of these signal molecules are released into the extracellular space by exocytosis from the signaling cell, as discussed in Chapter 13. Some, however, are emitted by diffusion through the signaling cell's plasma membrane, whereas others are displayed on the external surface of the cell and remain attached to it, providing a signal to other cells only when they make contact. Transmembrane proteins may be used for signaling in this way; or their extracellular domains may be released from the signaling cell's surface by proteolytic cleavage and then act at a distance.

Regardless of the nature of the signal, the *target cell* responds by means of a **receptor**, which specifically binds the signal molecule and then initiates a response in the target cell. The extracellular signal molecules often act at very low concentrations (typically $\leq 10^{-8}$ M), and the receptors that recognize them usually bind them with high affinity (affinity constant $K_a \geq 10^8$ liters/mole; see Figure 3–43).

In most cases, the receptors are transmembrane proteins on the target cell surface. When these proteins bind an extracellular signal molecule *(a ligand)*, they become activated and generate various intracellular signals that alter the behavior of the cell. In other cases, the receptor proteins are inside the target cell, and the signal molecule has to enter the cell to bind to them: this requires that the signal molecule be sufficiently small and hydrophobic to diffuse across the target cell's plasma membrane (**Figure 15–3**).

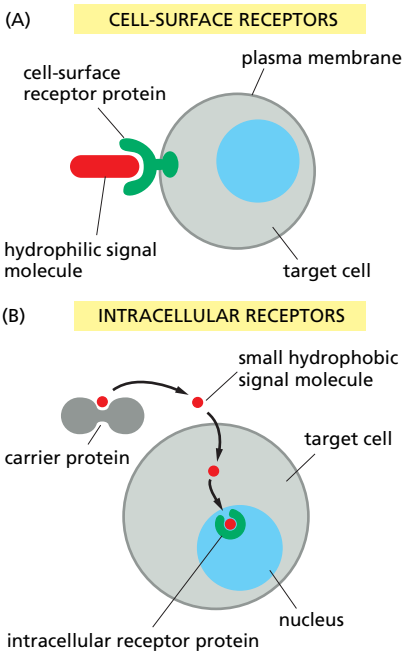


Figure 15–3 The binding of extracellular signal molecules to either cell-surface or intracellular receptors. (A) Most signal molecules are hydrophilic and are therefore unable to cross the target cell's plasma membrane directly; instead, they bind to cell-surface receptors, which in turn generate signals inside the target cell (see Figure 15–1). (B) Some small signal molecules, by contrast, diffuse across the plasma membrane and bind to receptor proteins inside the target cell—either in the cytosol or in the nucleus (as shown here). Many of these small signal molecules are hydrophobic and nearly insoluble in aqueous solutions; they are therefore transported in the bloodstream and other extracellular fluids bound to carrier proteins, from which they dissociate before entering the target cell.

Extracellular Signal Molecules Can Act Over Either Short or Long Distances

Many signal molecules remain bound to the surface of the signaling cell and influence only cells that contact it (**Figure 15–4**A). Such **contact-dependent signaling** is especially important during development and in immune responses. Contact-dependent signaling during development can sometimes operate over relatively large distances, where the communicating cells extend long processes to make contact with one another.

In most cases, however, signaling cells secrete signal molecules into the extracellular fluid. The secreted molecules may be carried far afield to act on distant target cells, or they may act as **local mediators**, affecting only cells in the local environment of the signaling cell. The latter process is called **paracrine signaling** (Figure 15–4B). Usually, the signaling and target cells in paracrine signaling are of different cell types, but cells may also produce signals that they themselves respond to: this is referred to as *autocrine signaling*. Cancer cells, for example, often use this strategy to stimulate their own survival and proliferation.

For paracrine signals to act only locally, the secreted molecules must not be allowed to diffuse too far; for this reason they are often rapidly taken up by neighboring target cells, destroyed by extracellular enzymes, or immobilized by the extracellular matrix. *Heparan sulfate proteoglycans* (discussed in Chapter 19), either in the extracellular matrix or attached to cell surfaces, often play a part in localizing the action of secreted signal proteins. They contain long polysaccharide side chains that bind the signal proteins and immobilize them. They may also control the stability of these proteins, their transport through the extracellular space, or their interaction with cell-surface receptors. Secreted protein *antagonists* also affect the distance over which a paracrine signal protein acts. These antagonists bind to either the signal molecule itself or its cell-surface receptor and block its activity, and they play an important part in restricting the effective range of secreted signal proteins that influence the developmental decisions that embryonic cells make (discussed in Chapter 22).

Large, complex, multicellular organisms need long-range signaling mechanisms to coordinate the behavior of cells in remote parts of the body. Thus, they have evolved cell types specialized for intercellular communication over large distances. The most sophisticated of these are nerve cells, or neurons, which typically extend long branching processes (axons) that enable them to contact target cells far away, where the processes terminate at the specialized sites of

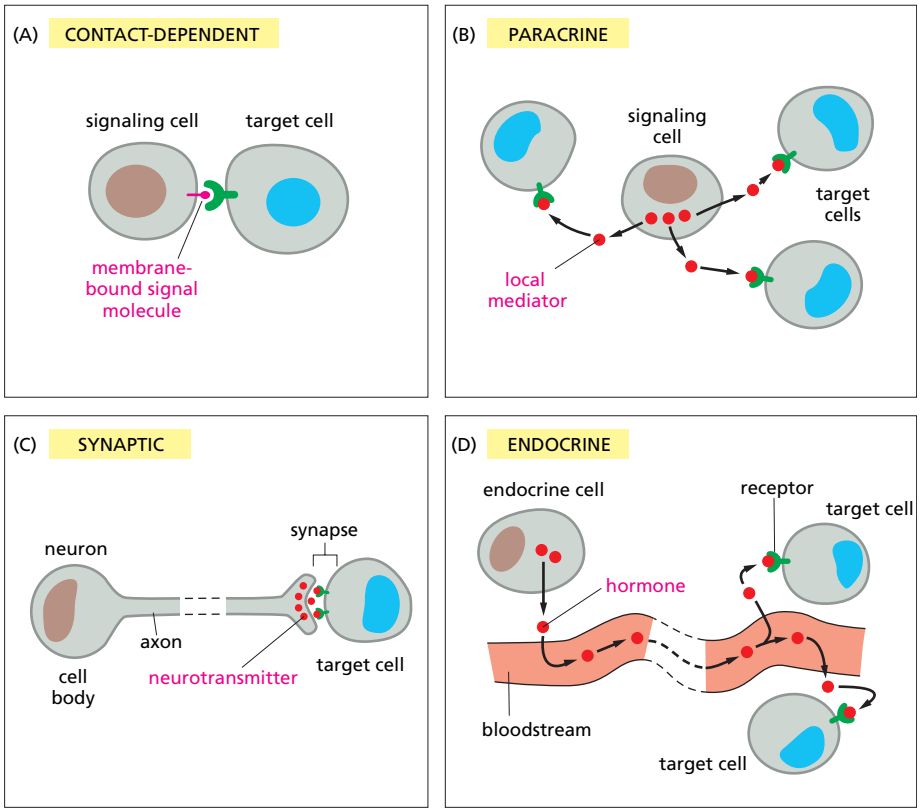


Figure 15–4 Four forms of intercellular signaling. (A) Contact-dependent signaling requires cells to be in direct membrane–membrane contact. (B) Paracrine signaling depends on signals that are released into the extracellular space and act locally on neighboring cells. (C) Synaptic signaling is performed by neurons that transmit signals electrically along their axons and release neurotransmitters at synapses, which are often located far away from the neuronal cell body. (D) Endocrine signaling depends on endocrine cells, which secrete hormones into the bloodstream for distribution throughout the body. Many of the same types of signaling molecules are used in paracrine, synaptic, and endocrine signaling; the crucial differences lie in the speed and selectivity with which the signals are delivered to their targets.

signal transmission known as *chemical synapses*. When activated by stimuli from the environment or from other nerve cells, the neuron sends electrical impulses (action potentials) rapidly along its axon; when such an impulse reaches the synapse at the end of the axon, it triggers secretion of a chemical signal that acts as a **neurotransmitter**. The tightly organized structure of the synapse ensures that the neurotransmitter is delivered specifically to receptors on the postsynaptic target cell (Figure 15–4C). The details of this **synaptic signaling** process are discussed in Chapter 11.

A quite different strategy for signaling over long distances makes use of **endocrine cells**. These secrete their signal molecules, called **hormones**, into the bloodstream, which carries the molecules far and wide, allowing them to act on target cells that may lie anywhere in the body (Figure 15–4D).

Figure 15–5 compares the mechanisms that allow endocrine cells and nerve cells to coordinate cell behavior over long distances in animals. Because endocrine signaling relies on diffusion and blood flow, it is relatively slow. Synaptic signaling, by contrast, is much faster, as well as more precise. Nerve cells can transmit information over long distances by electrical impulses that travel at speeds of up to 100 meters per second; once released from a nerve terminal, a neurotransmitter has to diffuse less than 100 nm to the target cell, a process that takes less than a millisecond. Another difference between endocrine and synaptic signaling is that, whereas hormones are greatly diluted in the bloodstream and interstitial fluid and therefore must be able to act at very low concentrations (typically $< 10^{-8}$ M), neurotransmitters are diluted much less and can achieve high local concentrations. The concentration of *acetylcholine* in the synaptic cleft of an active neuromuscular junction, for example, is about 5×10^{-4} M. Correspondingly, neurotransmitter receptors have a relatively low affinity for their ligand, which means that the neurotransmitter can dissociate rapidly from the receptor to help terminate a response. Moreover, after its release from a nerve terminal, a neurotransmitter is quickly removed from the synaptic cleft, either by specific hydrolytic enzymes that destroy it or by specific membrane transport proteins

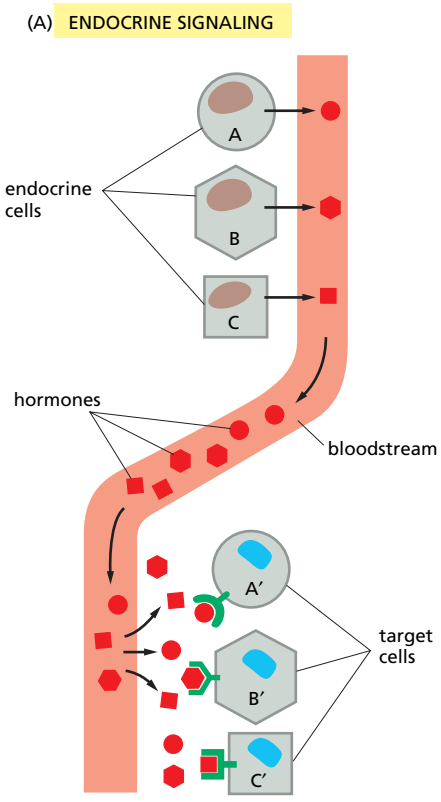


Figure 15–5 The contrast between endocrine and neuronal strategies for long-range signaling. In complex animals, endocrine cells and nerve cells work together to coordinate the activities of cells in widely separated parts of the body. Whereas different endocrine cells must use different hormones to communicate specifically with their target cells, different nerve cells can use the same neurotransmitter and still communicate in a highly specific manner. (A) Endocrine cells secrete hormones into the blood, and these act only on those target cells that carry the appropriate receptors: the receptors bind the specific hormone, which the target cells thereby "pull" from the extracellular fluid. (B) In synaptic signaling, by contrast, specificity arises from the synaptic contacts between a nerve cell and the specific target cells it signals. <CTGA> Usually, only a target cell that is in synaptic communication with a nerve cell is exposed to the neurotransmitter released from the nerve terminal (although some neurotransmitters act in a paracrine mode, serving as local mediators that influence multiple target cells in the area).

that pump it back into either the nerve terminal or neighboring glial cells. Thus, synaptic signaling is much more precise than endocrine signaling, both in time and in space.

The speed of a response to an extracellular signal depends not only on the mechanism of signal delivery but also on the nature of the target cell's response. When the response requires only changes in proteins already present in the cell, it can occur very rapidly: an allosteric change in a neurotransmitter-gated ion channel (discussed in Chapter 11), for example, can alter the plasma membrane electrical potential in milliseconds, and responses that depend solely on protein phosphorylation can occur within seconds. When the response involves changes in gene expression and the synthesis of new proteins, however, it usually requires many minutes or hours, regardless of the mode of signal delivery (**Figure 15–6**).

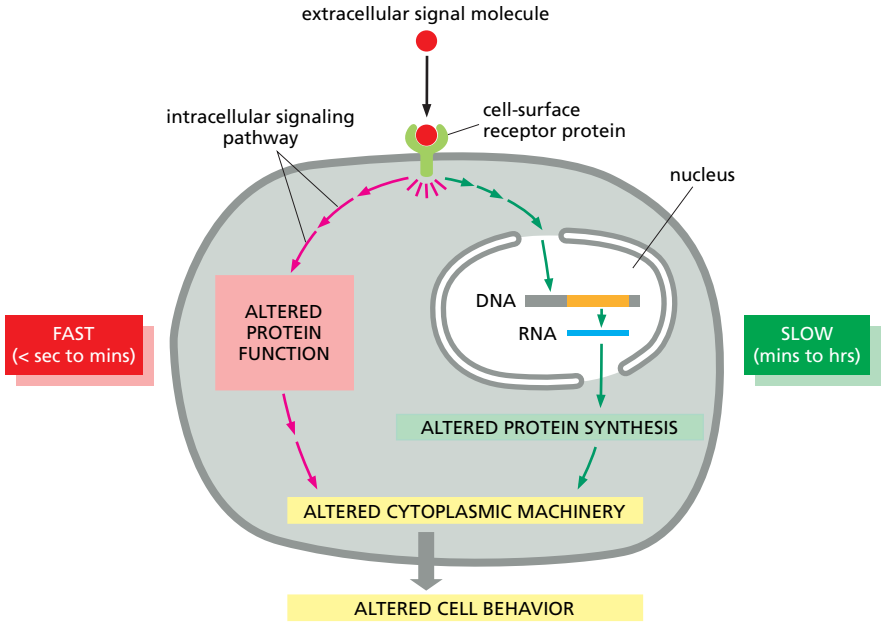


Figure 15–6 Extracellular signals can act slowly or rapidly to change the behavior of a target cell. Certain types of signaled responses, such as increased cell growth and division, involve changes in gene expression and the synthesis of new proteins; they therefore occur slowly, often starting after an hour or more. Other responses—such as changes in cell movement, secretion, or metabolism—need not involve changes in gene transcription and therefore occur much more quickly, often starting in seconds or minutes; they may involve the rapid phosphorylation of effector proteins in the cytoplasm, for example. Synaptic responses mediated by changes in membrane potential can occur in milliseconds (not shown).

Gap Junctions Allow Neighboring Cells to Share Signaling Information

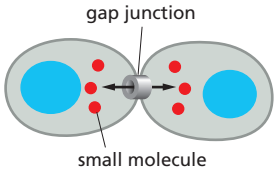


Figure 15–7 Signaling via gap junctions. Cells connected by gap junctions share small molecules, including small intracellular signaling molecules such as cyclic AMP and Ca^{2+}, and can therefore respond to extracellular signals in a coordinated way.

Gap junctions are narrow water-filled channels that directly connect the cytoplasms of adjacent epithelial cells, as well as of some other cell types (see Figure 19–34). The channels allow the exchange of inorganic ions and other small water-soluble molecules, but not of macromolecules such as proteins or nucleic acids. Thus, cells connected by gap junctions can communicate with each other directly, without having to surmount the barrier presented by the intervening plasma membranes (**Figure 15–7**). In this way, gap junctions provide for the most intimate of all forms of cell communication, short of cytoplasmic bridges (see Figure 21–31) or cell fusion.

In contrast with other modes of cell signaling, gap junctions generally allow communications to pass in both directions symmetrically, and their typical effect is to homogenize conditions in the communicating cells. They can also be important in spreading the effect of extracellular signals that act through small intracellular mediators such as Ca^{2+} and cyclic AMP (discussed later), which pass readily through the gap-junctional channels. In the liver, for example, a fall in blood glucose levels releases *noradrenaline* (norepinephrine) from sympathetic nerve endings. The noradrenaline stimulates hepatocytes in the liver to increase glycogen breakdown and to release glucose into the blood, a response that depends on an increase in intracellular cyclic AMP. Not all of the hepatocytes are innervated by sympathetic nerves, however. By means of the gap junctions that connect hepatocytes, the innervated hepatocytes transmit the signal to the noninnervated ones, at least in part by the movement of cyclic AMP through gap junctions. As expected, mice with a mutation in the major gap junction gene expressed in the liver fail to mobilize glucose normally when blood glucose levels fall. Gap junctions are discussed in detail in Chapter 19.

Each Cell Is Programmed to Respond to Specific Combinations of Extracellular Signal Molecules

A typical cell in a multicellular organism may be exposed to hundreds of different signal molecules in its environment. The molecules can be soluble, bound to the extracellular matrix, or bound to the surface of a neighboring cell; they can be stimulatory or inhibitory; they can act in innumerable different combinations; and they can influence almost any aspect of cell behavior. The cell must respond to this babel of signals selectively, according to its own specific character. This character is acquired through progressive cell specialization in the course of development. A cell may respond to one combination of signals by differentiating, to another combination by growing and dividing, and to yet another by performing some specialized function such as contraction or secretion. One of the great challenges in cell biology is to understand how a cell integrates all of this signaling information in order to make its crucial decisions—to divide, to move, to differentiate, and so on. For most of the cells in animal tissues, even the decision to continue living depends on correct interpretation of a specific combination of signals required for survival. When deprived of these signals (in a culture dish, for example), the cell activates a suicide program and kills itself—usually by *apoptosis,* a form of *programmed cell death* (**Figure 15–8**), as discussed in Chapter 18. Because different types of cells require different combinations of survival signals, each cell type is restricted to a specific set of environments in the body. Many epithelial cells, for example, require survival signals from the basal lamina on which they sit (discussed in Chapter 19); they die by apoptosis if they lose contact with this sheet of matrix.

In principle, the hundreds of signal molecules that an animal makes can be used in an almost unlimited number of signaling combinations so as to control the diverse behaviors of its cells in highly specific ways. Relatively small numbers of types of signal molecules and receptors are sufficient. The complexity lies in the ways in which cells respond to the combinations of signals that they receive.

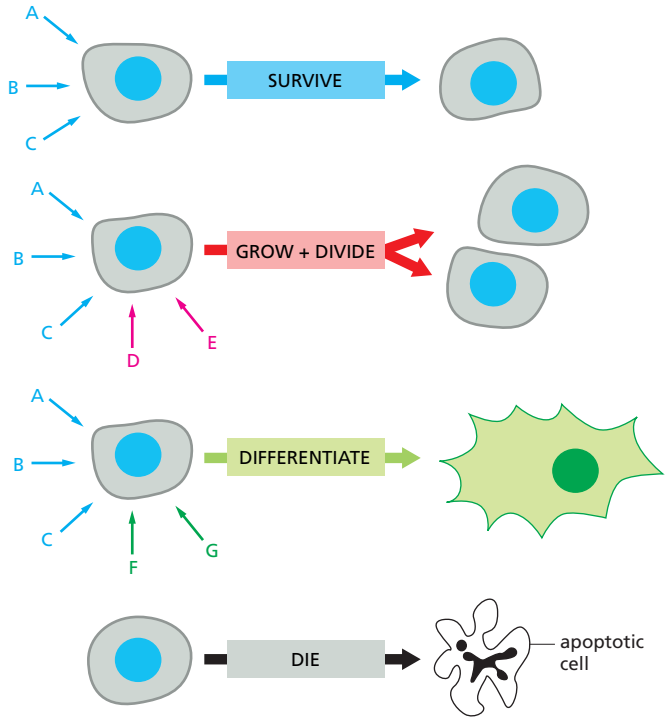


Figure 15–8 An animal cell's dependence on multiple extracellular signal molecules. Each cell type displays a set of receptors that enables it to respond to a corresponding set of signal molecules produced by other cells. These signal molecules work in combinations to regulate the behavior of the cell. As shown here, an individual cell often requires multiple signals to survive *(blue arrows)* and additional signals to grow and divide *(red arrow)* or differentiate *(green arrows)*. If deprived of appropriate survival signals, a cell will undergo a form of cell suicide known as apoptosis. The actual situation is even more complex. Although not shown, some extracellular signal molecules act to inhibit these and other cell behaviors, or even to induce apoptosis.

Different Types of Cells Usually Respond Differently to the Same Extracellular Signal Molecule

A cell's response to extracellular signals depends not only on the receptor proteins it possesses but also on the intracellular machinery by which it integrates and interprets the signals it receives. Thus, a single signal molecule usually has different effects on different types of target cells. The neurotransmitter acetylcholine (**Figure 15–9**A), for example, decreases the rate and force of contraction in heart muscle cells (Figure 15–9B), but it stimulates skeletal muscle cells to contract (see Figure 15–9C). In this case, the acetylcholine receptor proteins on skeletal muscle cells differ from those on heart muscle cells. But receptor

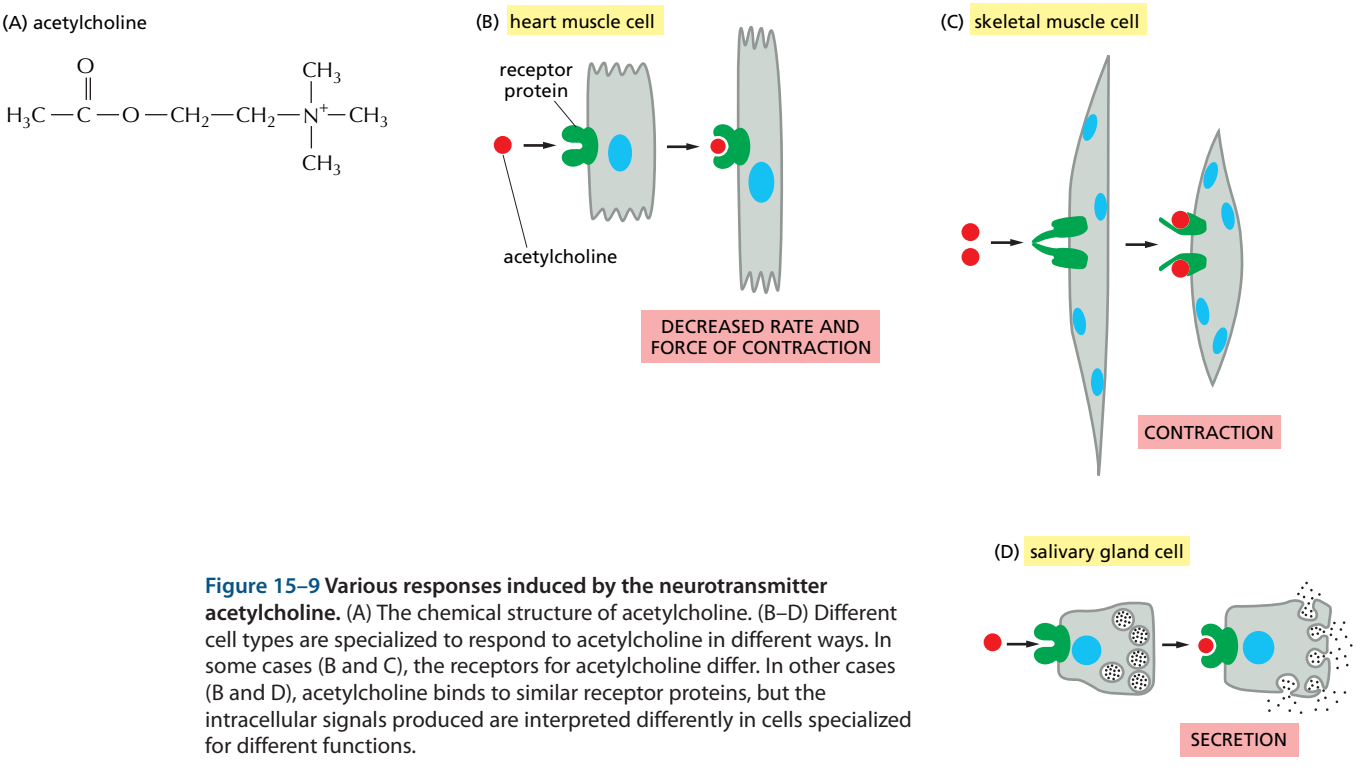


Figure 15–9 Various responses induced by the neurotransmitter acetylcholine. (A) The chemical structure of acetylcholine. (B–D) Different cell types are specialized to respond to acetylcholine in different ways. In some cases (B and C), the receptors for acetylcholine differ. In other cases (B and D), acetylcholine binds to similar receptor proteins, but the intracellular signals produced are interpreted differently in cells specialized for different functions.

differences are usually not the explanation for the different effects. The same signal molecule binding to identical receptor proteins usually produces very different responses in different types of target cells, as in the case of acetylcholine binding to heart muscle and to salivary gland cells (compare Figures 15–9B and D). In some cases, this reflects differences in the intracellular signaling proteins activated, whereas in others it reflects differences in the effector proteins or genes activated. Thus, an extracellular signal itself has little information content; it simply induces the cell to respond according to its predetermined state, which depends on the cell's developmental history and the specific genes it expresses.

The challenge of understanding how a cell integrates, processes, and reacts to the various inputs it receives is analogous in many ways to the challenge of understanding how the brain integrates and processes information to control behavior. In both cases, we need more than just a list of the components and connections in the system to understand how the process works. As a first step, we need to consider some basic principles concerning the way a cell responds to a simple signal of a given type.

The Fate of Some Developing Cells Depends on Their Position in Morphogen Gradients

The same signal acting on the same cell type can have qualitatively different effects depending on the signal's concentration. As we discuss in Chapter 22, such different responses of a cell to different concentrations of the same signal molecule are crucial in animal development, when cells are becoming different from one another.

The extracellular signal molecule in these cases during development is called a **morphogen**, and, in the simplest cases, it diffuses out from a localized cellular source (a *signaling center*), generating a signal concentration gradient. The responding cells adopt different cell fates in accordance with their position in the gradient: those cells closest to the signaling center that encounter the highest concentration of the morphogen have the highest number of receptors activated and follow one pathway of development, whereas those slightly further away follow another, and so on (**Figure 15–10**). As we discuss later (and in Chapter 22), the different levels of receptor activation lead to differences in the concentration or activity of one or more gene regulatory proteins in the nucleus of each cell, which in turn results in different patterns of gene expression. Further, more local signaling interactions between the cells in the gradient often help determine and stabilize the different fate choices.

A Cell Can Alter the Concentration of an Intracellular Molecule Quickly Only If the Lifetime of the Molecule Is Short

It is natural to think of signaling systems in terms of the changes produced when an extracellular signal is delivered. But it is just as important to consider what happens when the signal is withdrawn. During development, transient extracellular signals often produce lasting effects: they can trigger a change in the cell's development that persists indefinitely through cell memory mechanisms, as we discuss later (and in Chapters 7 and 22). In most cases in adult tissues, however, the response fades when a signal ceases. Often the effect is transient because the signal exerts its effects by altering the concentrations of intracellular molecules that are short-lived (unstable), undergoing continual turnover. Thus, once the extracellular signal is gone, the replacement of the old molecules by new ones wipes out all traces of the signal's action. It follows that the speed with which a cell responds to signal removal depends on the rate of destruction, or turnover, of the intracellular molecules that the signal affects.

It is also true, although much less obvious, that this turnover rate can determine the promptness of the response when an extracellular signal arrives. Consider, for example, two intracellular signaling molecules X and Y, both of which are normally maintained at a concentration of 1000 molecules per cell. The cell

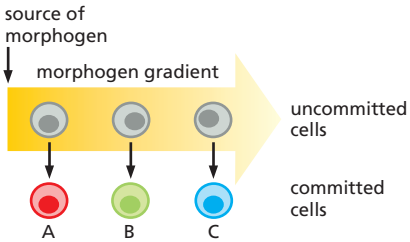


Figure 15–10 Cells adopt different fates depending on their position in a morphogen gradient. The different concentrations of morphogen induce the expression of different sets of genes, resulting in different cell fates (indicated by different letters and colors).

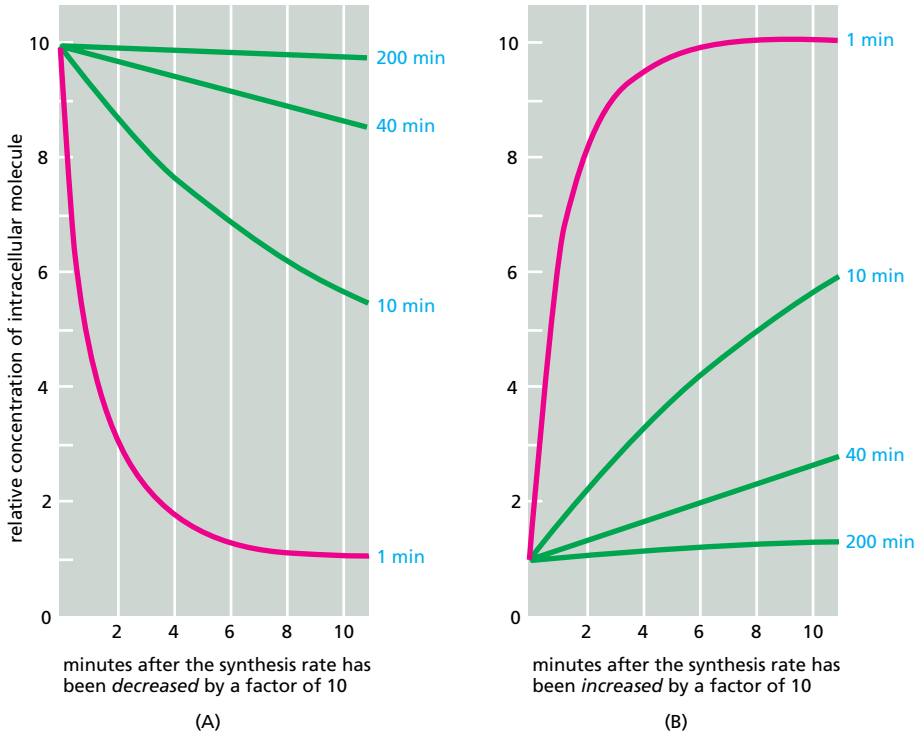


Figure 15–11 The importance of rapid turnover. The graphs show the predicted relative rates of change in the intracellular concentrations of molecules with differing turnover times when their synthesis rates are either (A) decreased or (B) increased suddenly by a factor of 10. In both cases, the concentrations of those molecules that are normally degraded rapidly in the cell *(red lines)* change quickly, whereas the concentrations of those that are normally degraded slowly *(green lines)* change proportionally more slowly. The numbers (in *blue)* on the right are the half-lives assumed for each of the different molecules.

synthesizes and degrades molecule Y at a rate of 100 molecules per second, with each molecule having an average lifetime of 10 seconds. Molecule X has a turnover rate that is 10 times slower than that of Y: it is both synthesized and degraded at a rate of 10 molecules per second, so that each molecule has an average lifetime in the cell of 100 seconds. If a signal acting on the cell causes a tenfold increase in the synthesis rates of both X and Y with no change in the molecular lifetimes, at the end of 1 second the concentration of Y will have increased by nearly 900 molecules per cell ($10 \times 100 - 100$), while the concentration of X will have increased by only 90 molecules per cell. In fact, after a molecule's synthesis rate has been either increased or decreased abruptly, the time required for the molecule to shift halfway from its old to its new equilibrium concentration is equal to its half-life—that is, equal to the time that would be required for its concentration to fall by half if all synthesis were stopped (**Figure 15–11**).

The same principles apply to proteins and small molecules, whether the molecules are in the extracellular space or inside cells. Many intracellular proteins have short half-lives, some surviving for less than 10 minutes. In most cases, these are key regulatory proteins whose concentrations are rapidly controlled in the cell by changes in their rates of synthesis.

We see later that many cell responses to extracellular signals depend on the conversion of intracellular signaling proteins from an inactive to an active form, rather than on their synthesis or degradation. Phosphorylation or the binding of GTP, for example, commonly activates signaling proteins. Even in these cases, however, the activation must be rapidly and continuously reversed (by dephosphorylation or GTP hydrolysis to GDP, respectively, in these examples) to make rapid signaling possible. These inactivation processes play a crucial part in determining the magnitude, rapidity, and duration of the response.

Nitric Oxide Gas Signals by Directly Regulating the Activity of Specific Proteins Inside the Target Cell

Most of this chapter is concerned with signaling pathways activated by cell-surface receptors. Before discussing these receptors and pathways, however, we briefly consider some important signal molecules that activate *intracellular receptors*. These molecules include nitric oxide and steroid hormones, which we discuss in turn. Although most extracellular signal molecules are hydrophilic

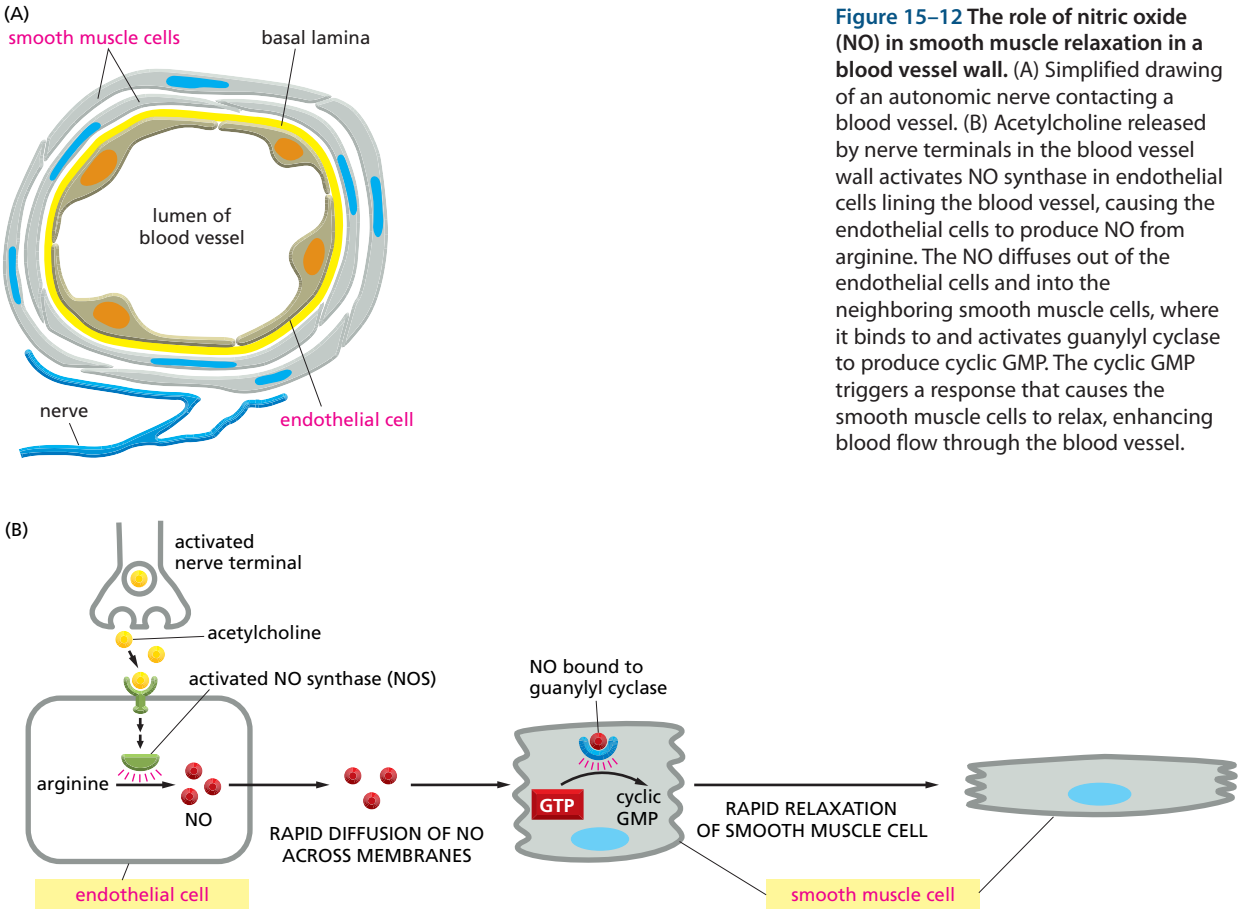


Figure 15–12 The role of nitric oxide (NO) in smooth muscle relaxation in a blood vessel wall. (A) Simplified drawing of an autonomic nerve contacting a blood vessel. (B) Acetylcholine released by nerve terminals in the blood vessel wall activates NO synthase in endothelial cells lining the blood vessel, causing the endothelial cells to produce NO from arginine. The NO diffuses out of the endothelial cells and into the neighboring smooth muscle cells, where it binds to and activates guanylyl cyclase to produce cyclic GMP. The cyclic GMP triggers a response that causes the smooth muscle cells to relax, enhancing blood flow through the blood vessel.

and bind to receptors on the surface of the target cell, some are hydrophobic enough, small enough, or both, to pass readily across the target cell's plasma membrane. Once inside, they directly regulate the activity of specific intracellular proteins. An important and remarkable example is the gas **nitric oxide (NO)**, which acts as a signal molecule in both animals and plants. Even some bacteria can detect a very low concentration of NO and move away from it.

In mammals, one of NO's many functions is to relax smooth muscle. It has this role in the walls of blood vessels, for example (**Figure 15–12**A). Autonomic nerves in the vessel wall release acetylcholine; the acetylcholine acts on the nearby endothelial cells that line the interior of the vessel; and the endothelial cells respond by releasing NO, which relaxes the smooth muscle cells in the wall, allowing the vessel to dilate. This effect of NO on blood vessels provides an explanation for the mechanism of action of nitroglycerine, which has been used for about 100 years to treat patients with angina (pain resulting from inadequate blood flow to the heart muscle). The nitroglycerine is converted to NO, which relaxes blood vessels. This reduces the workload on the heart and, as a consequence, reduces the oxygen requirement of the heart muscle.

Many types of nerve cells use NO more directly to signal to their neighbors. NO released by autonomic nerves in the penis, for example, causes the local blood vessel dilation that is responsible for penile erection. NO is also produced by activated macrophages and neutrophils to help them to kill invading microorganisms. In plants, NO is involved in the defensive responses to injury or infection.

NO is made by the deamination of the amino acid arginine, catalyzed by enzymes called **NO synthases** (**NOS**) (Figure 15–12B). The NOS in endothelial cells is called *eNOS,* while that in nerve and muscle cells is called *nNOS.* Nerve and muscle cells constitutively make nNOS, which is activated to produce NO by an influx of Ca^{2+} when the cells are stimulated. Macrophages, by contrast, make yet another NOS, called inducible NOS (*iNOS*) because they make it only when they are activated, usually in response to an infection.

Because dissolved NO passes readily across membranes, it rapidly diffuses out of the cell where it is produced and into neighboring cells (see Figure 15–12B). It acts only locally because it has a short half-life—about 5–10 seconds—in the extracellular space before oxygen and water convert it to nitrates and nitrites.

In some target cells, including smooth muscle cells, NO reversibly binds to iron in the active site of the enzyme *guanylyl cyclase,* stimulating this enzyme to produce the small intracellular signaling molecule *cyclic GMP,* which we discuss later. Thus, guanylyl cyclase acts both as an intracellular receptor for NO and as an intracellular signaling protein (see Figure 15–12B). NO can increase cyclic GMP in the cytosol within seconds, because the normal rate of turnover of cyclic GMP is high: a rapid degradation to GMP by a *phosphodiesterase* constantly balances the production of cyclic GMP from GTP by the guanylyl cyclase. The drug Viagra and its newer relatives inhibit the cyclic GMP phosphodiesterase in the penis, thereby increasing the amount of time that cyclic GMP levels remain elevated in the smooth muscle cells of penile blood vessels after NO production is induced by local nerve terminals. The cyclic GMP, in turn, keeps the blood vessels relaxed and thereby the penis erect.

NO can also signal cells independently of cyclic GMP. It can, for example, alter the activity of an intracellular protein by covalently nitrosylating thiol (–SH) groups on specific cysteines in the protein.

Carbon monoxide (CO) is another gas that is used as an extracellular signal molecule and like NO can act by stimulating guanylyl cyclase. But, these gases are not the only signal molecules that can pass directly across the target-cell plasma membrane, as we discuss next.

Nuclear Receptors Are Ligand-Modulated Gene Regulatory Proteins

Various small hydrophobic signal molecules diffuse directly across the plasma membrane of target cells and bind to intracellular receptors that are gene regulatory proteins. These signal molecules include steroid hormones, thyroid hormones, retinoids, and vitamin D. Although they differ greatly from one another in both chemical structure (**Figure 15–13**) and function, they all act by a similar mechanism. They bind to their respective intracellular receptor proteins and alter the ability of these proteins to control the transcription of specific genes. Thus, these proteins serve both as intracellular receptors and as intracellular effectors for the signal.

The receptors are all structurally related, being part of the very large **nuclear receptor superfamily**. Many family members have been identified by DNA sequencing only, and their ligand is not yet known; these proteins are therefore referred to as *orphan nuclear receptors.* Currently, more than half of the 48 nuclear receptors encoded in the human genome are orphans, as are 17 out of

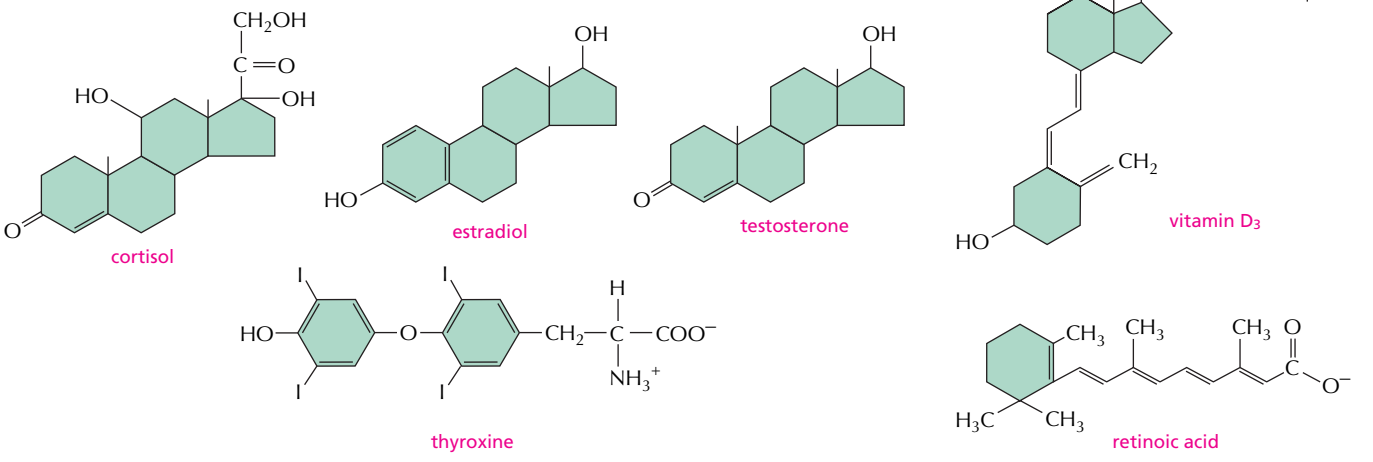


Figure 15–13 Some nongaseous signal molecules that bind to intracellular receptors. Note that all of them are small and hydrophobic. The active, hydroxylated form of vitamin D_3 is shown. Estradiol and testosterone are steroid sex hormones.

18 in *Drosophila* and all 278 in the nematode *C. elegans* (see Figure 4–85). Some mammalian nuclear receptors are regulated by intracellular metabolites rather than by secreted signal molecules; the *peroxisome proliferation-activated receptors (PPARs),* for example, bind intracellular lipid metabolites and regulate the transcription of genes involved in lipid metabolism and fat cell differentiation (discussed in Chapter 23). It seems likely that the nuclear receptors for hormones evolved from such receptors for intracellular metabolites, which would help explain their intracellular location.

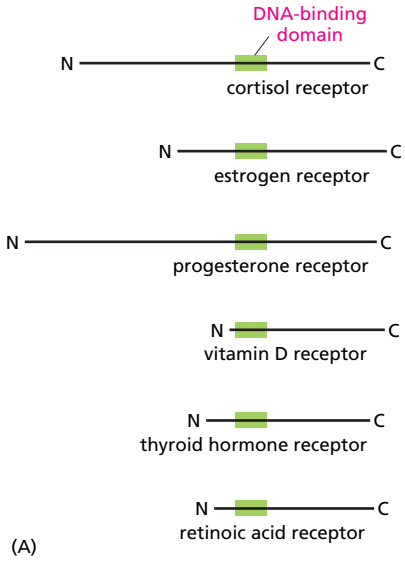
Steroid hormones—which include cortisol, the steroid sex hormones, vitamin D (in vertebrates), and the molting hormone *ecdysone* (in insects)—are all made from cholesterol. *Cortisol* is produced in the cortex of the adrenal glands and influences the metabolism of many types of cells. The *steroid sex hormones* are made in the testes and ovaries, and they are responsible for the secondary sex characteristics that distinguish males from females. *Vitamin D* is synthesized in the skin in response to sunlight; after it has been converted to its active form in the liver or kidneys, it regulates Ca^{2+} metabolism, promoting Ca^{2+} uptake in the gut and reducing its excretion in the kidneys. The *thyroid hormones,* which are made from the amino acid tyrosine, act to increase the metabolic rate of many cell types, while the *retinoids,* such as retinoic acid, are made from vitamin A and have important roles as local mediators in vertebrate development. Although all of these signal molecules are relatively insoluble in water, they are made soluble for transport in the bloodstream and other extracellular fluids by binding to specific carrier proteins, from which they dissociate before entering a target cell (see Figure 15–3B).

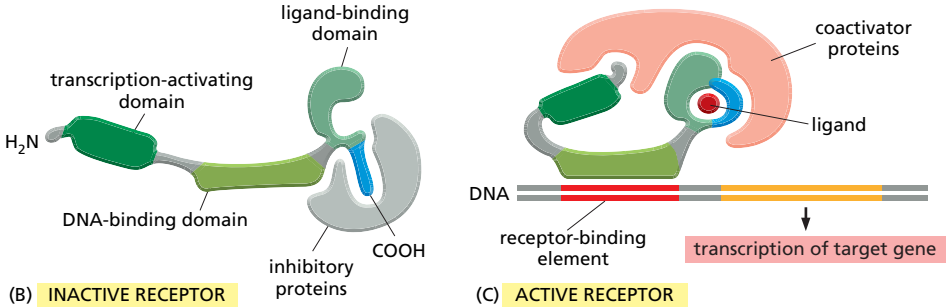
The nuclear receptors bind to specific DNA sequences adjacent to the genes that the ligand regulates. Some of the receptors, such as those for cortisol, are located primarily in the cytosol and enter the nucleus only after ligand binding; others, such as the thyroid and retinoid receptors, are bound to DNA in the nucleus even in the absence of ligand. In either case, the inactive receptors are usually bound to inhibitory protein complexes. Ligand binding alters the conformation of the receptor protein, causing the inhibitory complex to dissociate, while also causing the receptor to bind coactivator proteins that stimulate gene transcription (**Figure 15–14**). In other cases, however, ligand binding to a nuclear receptor inhibits transcription: some thyroid hormone receptors, for example, act as transcriptional activators in the absence of their hormone and become transcriptional repressors when hormone binds.

The transcriptional response usually takes place in multiple steps. In the cases in which ligand binding activates transcription, for example, the direct stimulation of a small number of specific genes occurs within about 30 minutes and constitutes the *primary response;* the protein products of these genes in turn activate other genes to produce a delayed, *secondary response;* and so on. In addition, some of the proteins produced in the primary response may act back to inhibit the transcription of primary response genes, thereby limiting the response—an example of negative feedback, which we discuss later. In this way, a simple hormonal trigger can cause a very complex change in the pattern of gene expression (**Figure 15–15**).

The responses to steroid and thyroid hormones, vitamin D, and retinoids are determined as much by the nature of the target cell as by the nature of the signal molecule. Many types of cells have the identical intracellular receptor, but the set of genes that the receptor regulates differs in each cell type. This is because more than one type of gene regulatory protein generally must bind to a eucaryotic gene to regulate its transcription. An intracellular receptor can therefore regulate a gene only if there is the right combination of other gene regulatory proteins, and many of these are cell-type specific.

In summary, each of these hydrophobic signal molecules induces a characteristic set of responses in an animal for two reasons. First, only certain types of cells have receptors for it. Second, each of these cell types contains a different combination of other cell-type-specific gene regulatory proteins that collaborate with the activated receptor to influence the transcription of specific sets of genes. This principle applies to all signaled responses that depend on gene regulatory proteins, including the many other examples we discuss in this chapter.





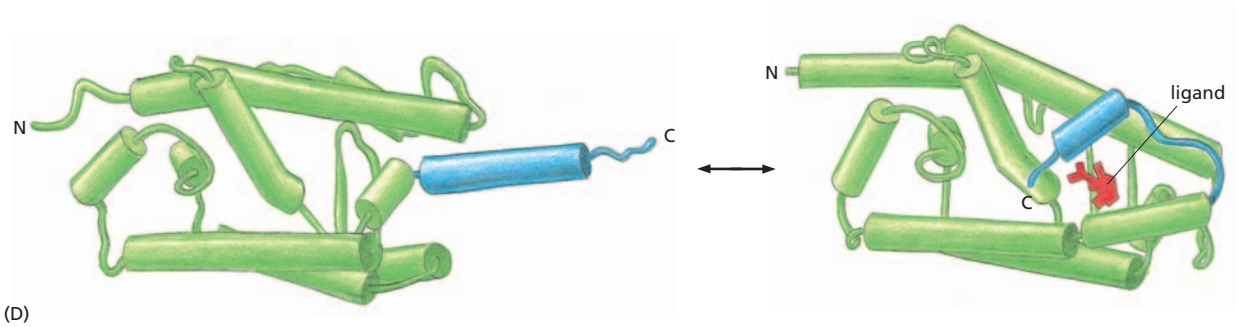


Figure 15–14 The nuclear receptor superfamily. All nuclear receptors bind to DNA as either homodimers or heterodimers, but for simplicity we show them as monomers here. (A) The receptors all have a related structure. Here, the short DNA-binding domain in each receptor is indicated in *light green.* (B) An inactive receptor protein is bound to inhibitory proteins. Domain-swap experiments suggest that many of the ligand-binding, transcription-activating, and DNA-binding domains in these receptors can function as interchangeable modules. (C) Receptor activation. Typically, the binding of ligand to the receptor causes the ligand-binding domain of the receptor to clamp shut around the ligand, the inhibitory proteins to dissociate, and coactivator proteins to bind to the receptor's transcription-activating domain, thereby increasing gene transcription. In other cases, ligand binding has the opposite effect, causing co-repressor proteins to bind to the receptor, thereby decreasing transcription (not shown). Though not shown here, activity can also be controlled through a change in the localization of the receptor: in its inactive form, it can be retained in the cytoplasm; ligand binding can then lead to the uncovering of nuclear localization signals that cause it to be imported into the nucleus to act on DNA. (D) The three-dimensional structure of a ligand-binding domain with *(right)* and without *(left)* ligand bound. Note that the *blue* α helix acts as a lid that snaps shut when the ligand (shown in *red*) binds, trapping the ligand in place.

The molecular details of how nuclear receptors and other gene regulatory proteins control specific gene transcription are discussed in Chapter 7.

Nuclear receptor proteins are sometimes also present on the cell surface, where they function by mechanisms entirely different from that just described. In the remainder of the chapter, we consider various ways in which cell-surface receptors convert extracellular signals into intracellular ones, a process called *signal transduction.*

The Three Largest Classes of Cell-Surface Receptor Proteins Are Ion-Channel-Coupled, G-Protein-Coupled, and Enzyme-Coupled Receptors

In contrast to the small hydrophobic signal molecules just discussed that bind to intracellular receptors, most extracellular signal molecules bind to specific receptor proteins on the surface of the target cells they influence and do not enter the cytosol or nucleus. These cell-surface receptors act as *signal transducers* by converting an extracellular ligand-binding event into intracellular signals that alter the behavior of the target cell.

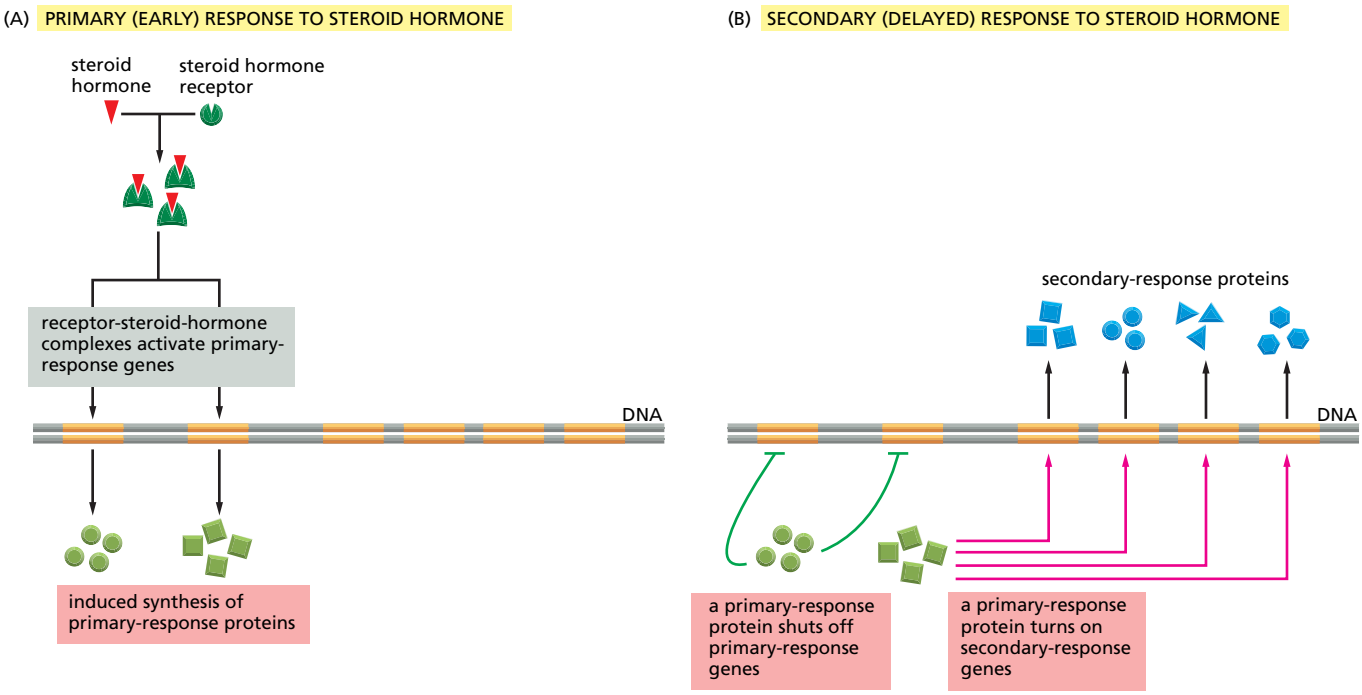


Figure 15–15 An example of primary and secondary responses induced by the activation of a nuclear hormone receptor. (A) Early primary response and (B) delayed secondary response. Some of the primary-response proteins turn on secondary-response genes, while others turn off the primary-response genes. The actual number of primary-response and secondary-response genes is greater than shown. Drugs that inhibit protein synthesis suppress the transcription of secondary-response genes but not primary-response genes, allowing these two classes of gene transcription responses to be readily distinguished. The figure shows the responses to a steroid hormone, but the same principles apply for many other ligands that activate nuclear receptor proteins.

Most cell-surface receptor proteins belong to one of three classes, defined by their transduction mechanism. **Ion-channel-coupled receptors**, also known as *transmitter-gated ion channels* or *ionotropic receptors,* are involved in rapid synaptic signaling between nerve cells and other electrically excitable target cells such as nerve and muscle cells (**Figure 15–16**A). This type of signaling is mediated by a small number of neurotransmitters that transiently open or close an ion channel formed by the protein to which they bind, briefly changing the ion permeability of the plasma membrane and thereby the excitability of the postsynaptic target cell. Most ion-channel-coupled receptors belong to a large family of homologous, multipass transmembrane proteins. Because they are discussed in detail in Chapter 11, we shall not consider them further here.

G-protein-coupled receptors act by indirectly regulating the activity of a separate plasma-membrane-bound target protein, which is generally either an enzyme or an ion channel. A *trimeric GTP-binding protein (G protein)* mediates the interaction between the activated receptor and this target protein (Figure 15–16B). The activation of the target protein can change the concentration of one or more small intracellular mediators (if the target protein is an enzyme), or it can change the ion permeability of the plasma membrane (if the target protein is an ion channel). The small intracellular mediators act in turn to alter the behavior of yet other signaling proteins in the cell. All of the G-protein-coupled receptors belong to a large family of homologous, multipass transmembrane proteins.

Enzyme-coupled receptors either function directly as enzymes or associate directly with enzymes that they activate (Figure 15–16C). They are usually single-pass transmembrane proteins that have their ligand-binding site outside the cell and their catalytic or enzyme-binding site inside. Enzyme-coupled receptors are heterogeneous in structure compared with the other two classes. The great majority, however, are either protein kinases or associate with protein kinases, which phosphorylate specific sets of proteins in the target cell when activated.

There are also some types of cell-surface receptors that do not fit easily into any of these classes but have important functions in controlling the specialization of different cell types during development and in tissue renewal and repair. We discuss these in a later section, after we explain how G-protein-coupled receptors and enzyme-coupled receptors operate. First, however, we consider some general principles of signaling via cell-surface receptors, in order to prepare for the detailed discussion of the major classes of cell-surface receptors that follows.

Most Activated Cell-Surface Receptors Relay Signals Via Small Molecules and a Network of Intracellular Signaling Proteins

A combination of small and large *intracellular signaling molecules* relays signals received at the cell surface by either G-protein-coupled or enzyme-coupled receptors into the cell interior. The resulting chain of intracellular signaling events ultimately alters effector proteins that are responsible for modifying the behavior of the cell (see Figure 15–1).

The small intracellular signaling molecules are called **small intracellular mediators**, or **second messengers** (the "first messengers" being the extracellular signals). They are generated in large numbers in response to receptor activation and often diffuse away from their source, spreading the signal to other parts of the cell. Some, such as *cyclic AMP* and Ca^{2+}, are water-soluble and diffuse in the cytosol, while others, such as *diacylglycerol,* are lipid-soluble and diffuse in the plane of the plasma membrane. In either case, they pass the signal on by binding to and altering the conformation and behavior of selected signaling proteins or effector proteins.

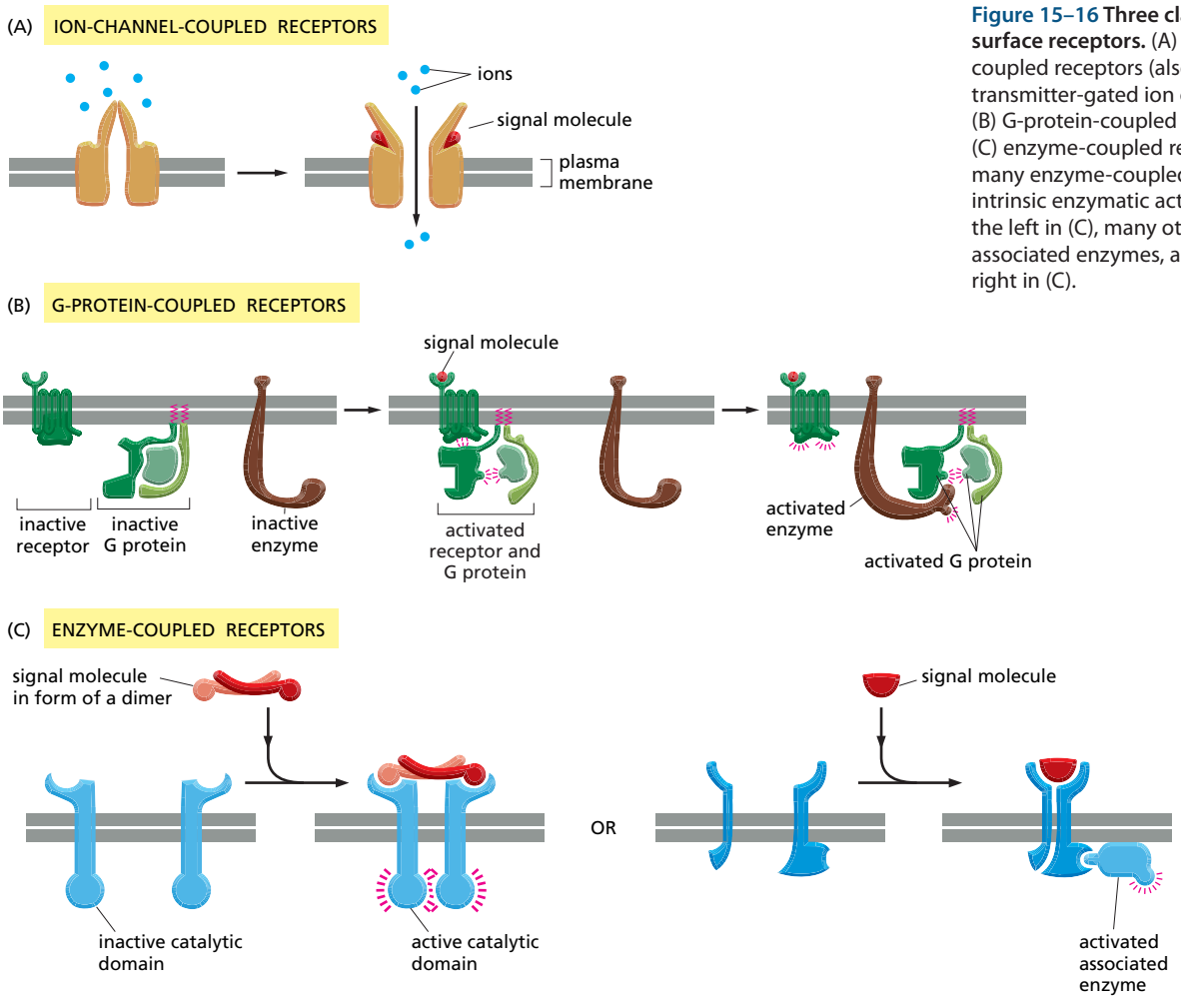


Figure 15–16 Three classes of cell-surface receptors. (A) Ion-channel-coupled receptors (also called transmitter-gated ion channels), (B) G-protein-coupled receptors, and (C) enzyme-coupled receptors. Although many enzyme-coupled receptors have intrinsic enzymatic activity, as shown on the left in (C), many others rely on associated enzymes, as shown on the right in (C).

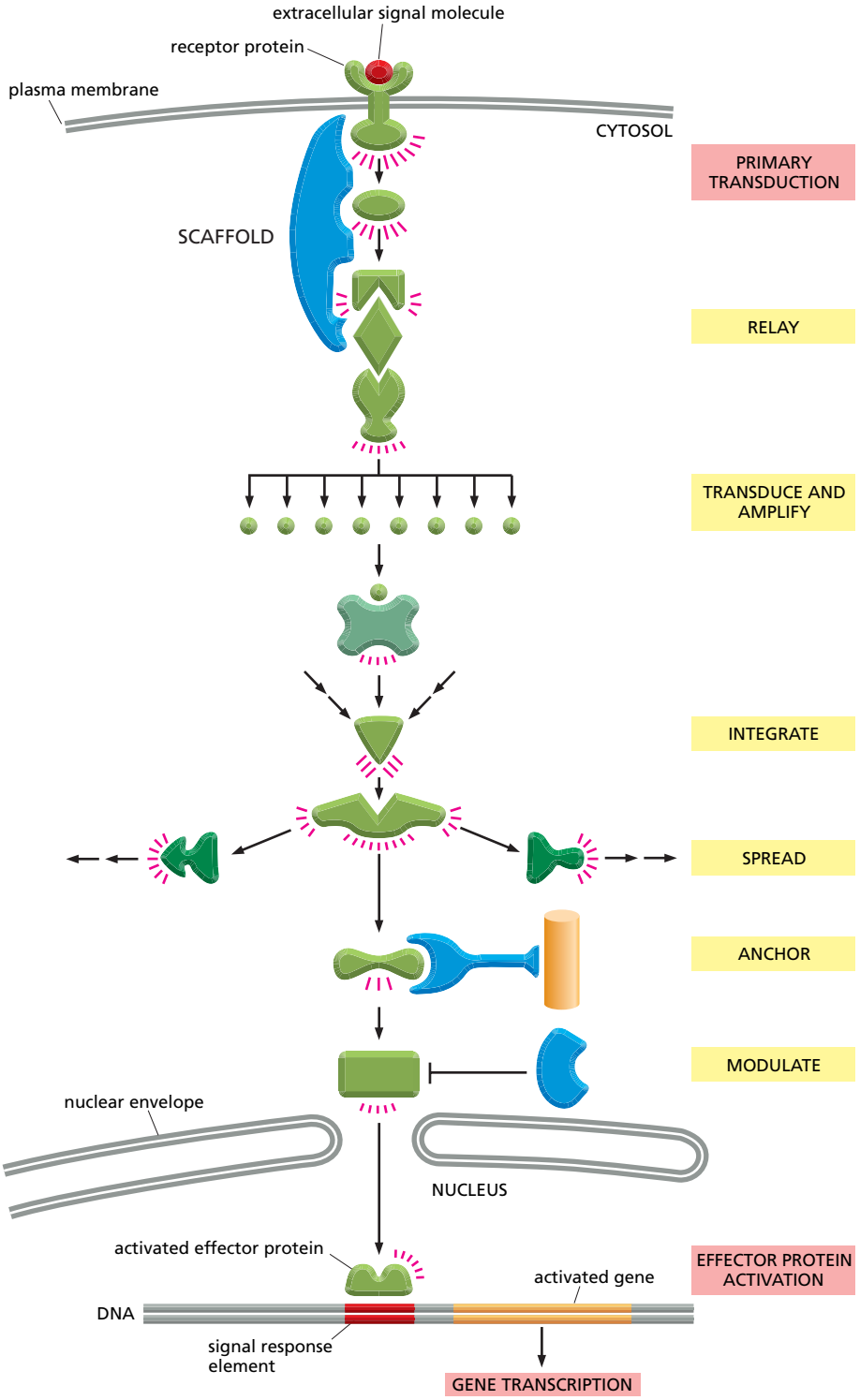


Figure 15–17 A hypothetical intracellular signaling pathway from a cell-surface receptor to the nucleus. In this example, a series of signaling proteins and small intracellular mediators relay the extracellular signal into the nucleus, causing a change in gene expression. The signal is altered (transduced), amplified, distributed, and modulated *en route*. Because many of the steps can be affected by other extracellular and intracellular signals, the final effect of one extracellular signal depends on multiple factors affecting the cell (see Figure 15–8). Ultimately, the signaling pathway activates (or inactivates) effector proteins that alter the cell's behavior. In this example, the effector is a gene regulatory protein that activates gene transcription. Although the figure shows individual signaling proteins performing a single function, in actuality they often have more than one function; scaffold proteins, for example, often also serve to anchor several signaling proteins to a particular intracellular structure.

Most signaling pathways to the nucleus are more direct than this one, which is not based on a known pathway.

The large intracellular signaling molecules are **intracellular signaling proteins**, which help relay the signal into the cell by either generating small intracellular mediators or activating the next signaling or effector protein in the pathway. These proteins form a functional network, in which each protein helps to process the signal in one or more of the following ways as it spreads the signal's influence through the cell (**Figure 15–17**):

1. The protein may simply *relay* the signal to the next signaling component in the chain.
2. It may act as a *scaffold* to bring two or more signaling proteins together so that they can interact more quickly and efficiently.

3. It may transform, or *transduce,* the signal into a different form, which is suitable for either passing the signal along or stimulating a cell response.
4. It may *amplify* the signal it receives, either by producing large amounts of a small intracellular mediator or by activating many copies of a downstream signaling protein. In this way, a small number of extracellular signal molecules can evoke a large intracellular response. When there are multiple amplification steps in a relay chain, the chain is often referred to as a **signaling cascade**.
5. It may receive signals from two or more signaling pathways and *integrate* them before relaying a signal onward. A protein that requires input from two or more signaling pathways to become activated, is often referred to as a coincidence detector.
6. It may *spread* the signal from one signaling pathway to another, creating branches in the signaling stream, thereby increasing the complexity of the response.
7. It may *anchor* one or more signaling proteins in a pathway to a particular structure in the cell where the signaling proteins are needed.
8. It may *modulate* the activity of other signaling proteins and thereby regulate the strength of signaling along a pathway.

We now consider in more detail some of the strategies that intracellular signaling proteins use in processing the signal as it is relayed along signaling pathways. We encounter these general strategies again in later sections of the chapter, when we discuss specific receptor classes and the signaling pathways they activate.

Many Intracellular Signaling Proteins Function as Molecular Switches That Are Activated by Phosphorylation or GTP Binding

Many intracellular signaling proteins behave like *molecular switches.* When they receive a signal, they switch from an inactive to an active conformation, until another process switches them off, returning them to their inactive conformation. As we mentioned earlier, the switching off is just as important as the switching on. If a signaling pathway is to recover after transmitting a signal so that it can be ready to transmit another, every activated molecule in the pathway must return to its original, unactivated state.

Two important classes of molecular switches that operate in intracellular signaling pathways depend on the gain or loss of phosphate groups for their activation or inactivation, although the way in which the phosphate is gained and lost is very different in the two classes. The largest class consists of proteins that are activated or inactivated by **phosphorylation** (discussed in Chapter 3). For these proteins, the switch is thrown in one direction by a **protein kinase**, which covalently adds one or more phosphate groups to the signaling protein, and in the other direction by a **protein phosphatase**, which removes the phosphate groups (**Figure 15–18**A). The activity of any protein regulated by

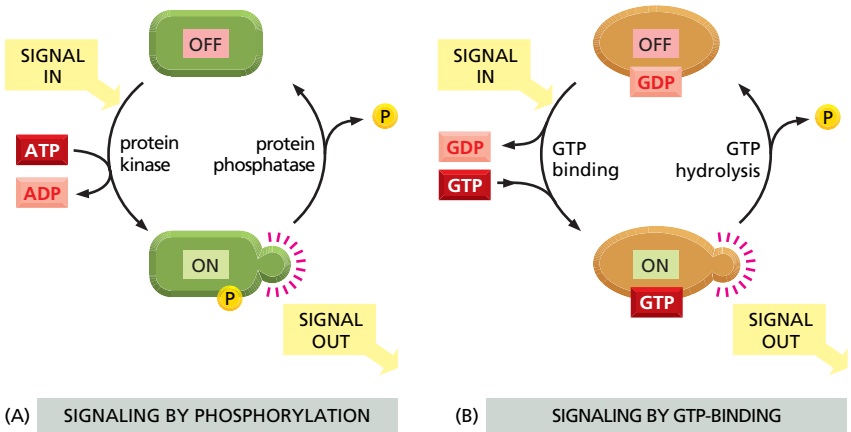


Figure 15–18 Two types of intracellular signaling proteins that act as molecular switches. Although one type is activated by phosphorylation and the other by GTP binding, in both cases the addition of a phosphate group switches the activation state of the protein and the removal of the phosphate switches it back again. (A) A protein kinase covalently adds a phosphate from ATP to the signaling protein, and a protein phosphatase removes the phosphate. Although not shown, some signaling proteins are activated by dephosphorylation rather than by phosphorylation. (B) A GTP-binding protein is induced to exchange its bound GDP for GTP, which activates the protein; the protein then inactivates itself by hydrolyzing its bound GTP to GDP.

phosphorylation depends on the balance at any instant between the activities of the kinases that phosphorylate it and of the phosphatases that dephosphorylate it. About 30% of human proteins contain covalently attached phosphate, and the human genome encodes about 520 protein kinases and about 150 protein phosphatases. It is thought that a typical mammalian cell makes use of hundreds of distinct types of protein kinases at any one time.

Many signaling proteins controlled by phosphorylation are themselves protein kinases, and these are often organized into **phosphorylation cascades**. In such a cascade, one protein kinase, activated by phosphorylation, phosphorylates the next protein kinase in the sequence, and so on, relaying the signal onward and, in the process, amplifying it and sometimes spreading it to other signaling pathways. Two main types of protein kinases operate as intracellular signaling proteins. The great majority are **serine/threonine kinases**, which phosphorylate proteins on serines and (less often) threonines. Others are **tyrosine kinases**, which phosphorylate proteins on tyrosines. An occasional kinase can do both.

The other important class of molecular switches that function by gaining and losing phosphate groups consists of **GTP-binding proteins** (discussed in Chapter 3). These proteins switch between an "on" (actively signaling) state when GTP is bound and an "off" state when GDP is bound. In the "on" state, they have intrinsic GTPase activity and shut themselves off by hydrolyzing their bound GTP to GDP (Figure 15–18B). There are two major types of GTP-binding proteins. Large *trimeric GTP-binding proteins* (also called *G proteins*) help relay signals from G-protein-coupled receptors that activate them (see Figure 15–16B). Small **monomeric GTPases** (also called *monomeric GTP-binding proteins*) help relay signals from many classes of cell-surface receptors.

Specific regulatory proteins control both types of GTP-binding proteins. **GTPase-activating proteins** (**GAPs**) drive the proteins into an "off" state by increasing the rate of hydrolysis of bound GTP; the GAPs that function in this way are also called *regulators of G-protein signaling (RGS)* proteins. Conversely, G-protein-coupled receptors activate trimeric G proteins, and **guanine nucleotide exchange factors** (**GEFs**) activate monomeric GTPases, by promoting the release of bound GDP in exchange for binding of GTP. **Figure 15–19** illustrates the regulation of monomeric GTPases.

Both trimeric G proteins and monomeric GTPases also participate in many other processes in eucaryotic cells, including the regulation of vesicular traffic and aspects of cell division.

As discussed earlier, specific combinations of extracellular signals, rather than one signal molecule acting alone, are generally required to stimulate complex cell behaviors, such as cell survival and cell growth and proliferation (see Figure 15–8). The cell therefore has to integrate information coming from multiple signals if it is to make an appropriate response; many mammalian cells, for example, require both soluble signals and signals from the extracellular matrix

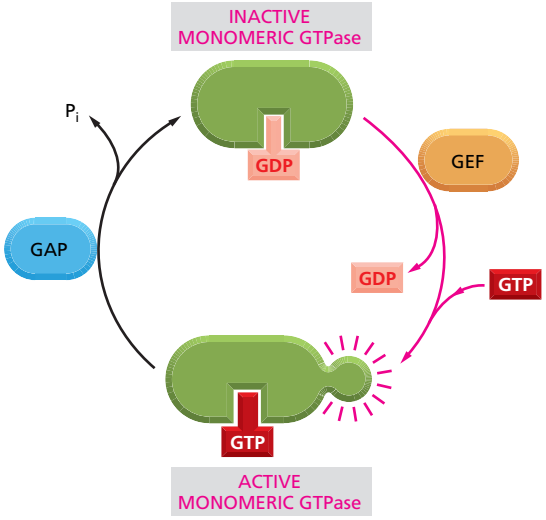


Figure 15–19 The regulation of a monomeric GTPase. GTPase-activating proteins (GAPs) inactivate the protein by stimulating it to hydrolyze its bound GTP to GDP, which remains tightly bound to the inactivated GTPase. Guanine nucleotide exchange factors (GEFs) activate the inactive protein by stimulating it to release its GDP; because the concentration of GTP in the cytosol is 10 times greater than the concentration of GDP, the protein rapidly binds GTP once it ejects GDP and is thereby activated.

(discussed in Chapter 19) to grow and proliferate. The integration depends in part on intracellular *coincidence detectors*, which are equivalent to *AND gates* in the microprocessor of a computer, in that they are only activated if they receive multiple converging signals (see Figure 15–17). **Figure 15–20** illustrates a simple hypothetical example of such a protein.

Not all molecular switches in signaling pathways depend on phosphorylation or GTP binding, however. We see later that some signaling proteins are switched on or off by the binding of another signaling protein or a small intracellular mediator such as cyclic AMP or Ca^{2+}, or by covalent modifications other than phosphorylation or dephosphorylation, such as ubiquitylation. Moreover, not all intracellular signaling proteins act as switches when they are phosphorylated or otherwise reversibly modified. As we discuss later, in many cases the covalently added group simply marks the protein so that it can interact with other signaling proteins that recognize the modification.

Intracellular Signaling Complexes Enhance the Speed, Efficiency, and Specificity of the Response

Even a single type of extracellular signal acting through a single type of cell-surface receptor often activates multiple parallel signaling pathways and can thereby influence multiple aspects of cell behavior—such as shape, movement, metabolism, and gene expression. Given the complexity of signal-response systems, which often involve multiple interacting relay chains of signaling proteins, how does an individual cell manage to make specific responses to so many different combinations of extracellular signals? The question is especially puzzling because many of the signals are closely related to one another and bind to closely related types of receptors. The same type of intracellular relay protein may couple one receptor subtype to one set of effectors and another receptor subtype to another set of effectors. In such cases, how is it possible to achieve specificity and avoid cross-talk? One strategy makes use of **scaffold proteins** (see Figure 15–17), which bind together groups of interacting signaling proteins into *signaling complexes*, often before a signal has been received (**Figure 15–21**A). Because the scaffold holds the signaling proteins in close proximity, the components can interact at high local concentrations and be sequentially activated speedily, efficiently, and selectively in response to an appropriate extracellular signal, avoiding unwanted cross-talk with other signaling pathways.

In other cases, signaling complexes form only transiently in response to an extracellular signal and rapidly disassemble when the signal is gone. Such transient complexes often assemble around a receptor after an extracellular signal molecule has activated it. In many of these cases, the cytoplasmic tail of the activated receptor is phosphorylated during the activation process, and the phosphorylated amino acids then serve as docking sites for the assembly of other signaling proteins (Figure 15–21B). In yet other cases, receptor activation leads to the production of modified phospholipid molecules (called phosphoinositides) in the adjacent plasma membrane, which then recruit specific intracellular signaling proteins to this region of membrane, where they are activated (Figure 15–21C). We discuss the roles of phosphoinositides in membrane trafficking events in Chapter 13.

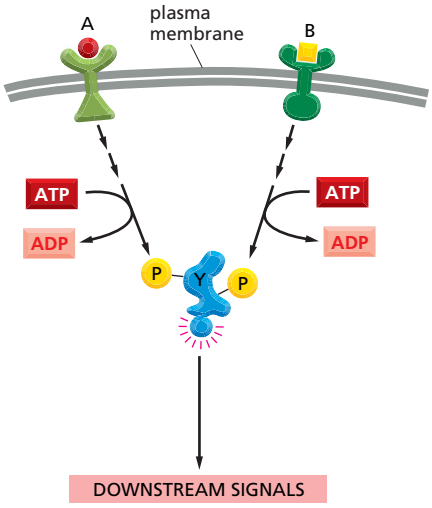


Figure 15–20 Signal integration. Extracellular signals A and B activate different intracellular signaling pathways, each of which leads to the phosphorylation of protein Y but at different sites on the protein. Protein Y is activated only when both of these sites are phosphorylated, and therefore it becomes active only when signals A and B are simultaneously present. Such proteins are often called coincidence detectors.

Modular Interaction Domains Mediate Interactions Between Intracellular Signaling Proteins

Simply bringing intracellular signaling proteins together into close proximity is sometimes sufficient to activate them. Thus, *induced proximity*, where a signal triggers assembly of a signaling complex, is commonly used to relay signals from protein to protein along a signaling pathway. The assembly of such signaling complexes depends on various highly conserved, small **interaction domains,** which are found in many intracellular signaling proteins. Each of these compact

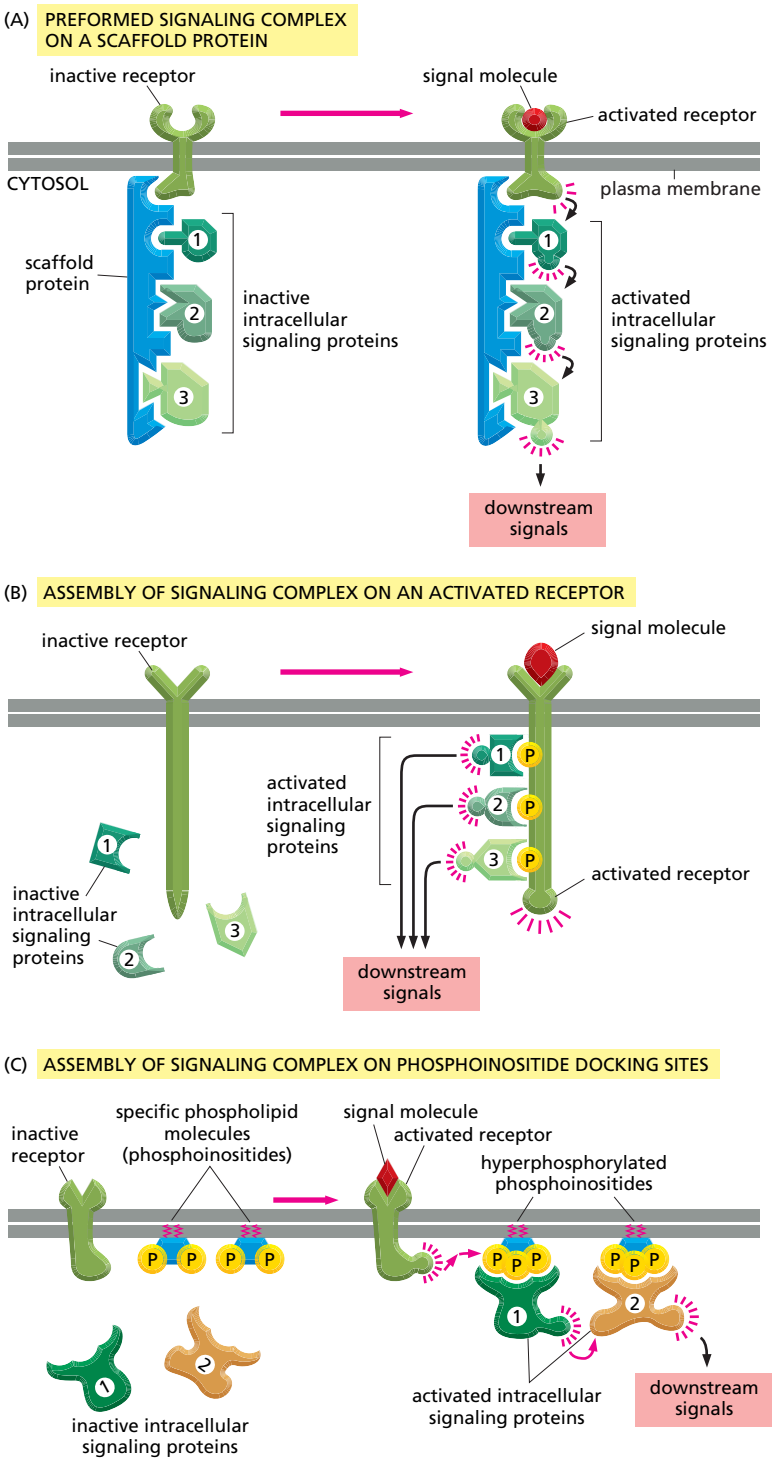


Figure 15–21 Three types of intracellular signaling complexes. (A) A receptor and some of the intracellular signaling proteins it activates in sequence are preassembled into a signaling complex on the inactive receptor by a large scaffold protein. In other cases, a preassembled complex binds to the receptor only after the receptor is activated (not shown). (B) A signaling complex assembles on a receptor only after the binding of an extracellular signal molecule has activated the receptor; here the activated receptor phosphorylates itself at multiple sites, which then act as docking sites for intracellular signaling proteins. (C) Activation of a receptor leads to the increased phosphorylation of specific phospholipids (phosphoinositides) in the adjacent plasma membrane, which then serve as docking sites for specific intracellular signaling proteins, which can now interact with each other.

protein modules binds to a particular structural motif in another protein (or lipid) molecule with which the signaling protein interacts. The recognized motif in the interacting protein can be a short peptide sequence, a covalent modification (such as phosphorylated or ubiquitylated amino acids), or another protein domain. The use of modular interaction domains presumably facilitated the evolution of new signaling pathways; because it can be inserted almost anywhere in a protein without disturbing the protein's folding or function, a new interaction domain added to an existing signaling protein could connect the protein to additional signaling pathways.

There are many types of interaction domains in signaling proteins. *Src homology 2 (SH2) domains* and *phosphotyrosine-binding (PTB) domains*, for example, bind to phosphorylated tyrosines in a particular peptide sequence on

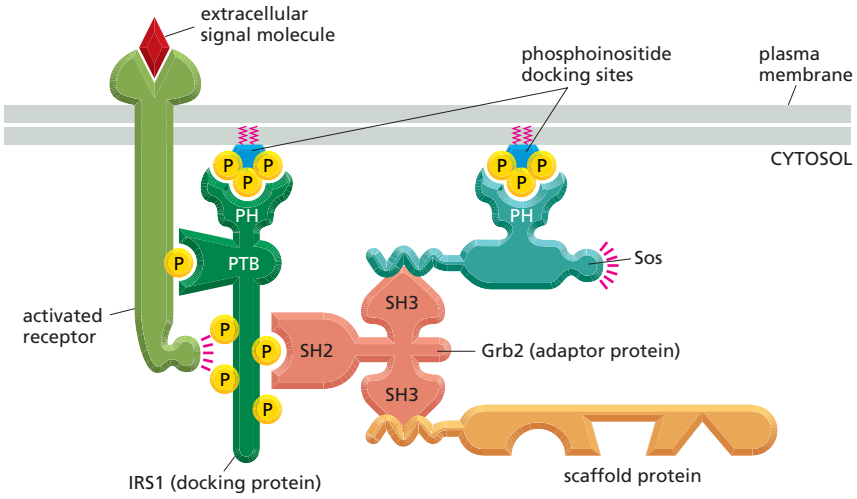


Figure 15–22 A specific signaling complex formed using modular interaction domains. This example is based on the insulin receptor, which is an enzyme-coupled receptor (a receptor tyrosine kinase, discussed later). First, the activated receptor phosphorylates itself on tyrosines, and one of the phosphotyrosines then recruits a docking protein called insulin receptor substrate-1 (IRS1) via a PTB domain of IRS1; the PH domain of IRS1 also binds to specific phosphoinositides on the inner surface of the plasma membrane. Then, the activated receptor phosphorylates IRS1 on tyrosines, and one of these phosphotyrosines recruits the adaptor protein (Grb2) via an SH2 domain of Grb2. Next, Grb2 uses one of its two SH3 domains to bind to a proline-rich region of the monomeric GTPase-activating protein called Sos (a Ras-GEF, discussed later), which also binds to phosphoinositides in the plasma membrane via its PH domain. Grb2 uses its other SH3 domain to bind to a proline-rich sequence in a scaffold protein. The scaffold protein binds several other signaling proteins, and the other phosphorylated tyrosines on IRS1 recruit additional signaling proteins that have SH2 domains (not shown).

activated receptors or intracellular signaling proteins. *Src homology 3 (SH3)* domains bind to short proline-rich amino acid sequences. Some *pleckstrin homology (PH)* domains bind to the charged head groups of specific phosphoinositides that are produced in the plasma membrane in response to an extracellular signal; they enable the protein they are part of to dock on the membrane and interact with other similarly recruited signaling proteins (see Figure 15–21C). Some signaling proteins consist solely of two or more interaction domains and function only as **adaptors** to link two other proteins together in a signaling pathway.

Interaction domains enable signaling proteins to bind to one another in multiple specific combinations. Like Lego bricks, the proteins can form linear or branching chains or three-dimensional networks, which determine the route followed by the signaling pathway. As an example, **Figure 15–22** illustrates how some interaction domains function in the case of the receptor for the hormone *insulin*.

Some cell-surface receptors and intracellular signaling proteins may cluster together transiently in specific microdomains in the lipid bilayer of the plasma membrane that are enriched in cholesterol and glycolipids (see Figure 10–13). These *lipid rafts* may promote efficient signaling by serving as sites where signaling molecules assemble and interact, but their importance in signaling remains controversial.

Another way of bringing receptors and intracellular signaling proteins together is to concentrate them in a specific region of the cell. An important example is the **primary cilium** that projects like an antenna from the surface of most vertebrate cells (discussed in Chapter 16). It is usually short and nonmotile and has microtubules in its core, and a number of surface receptors and signaling proteins are concentrated there. We shall see later that light and smell receptors are also highly concentrated in specialized cilia.

Cells Can Use Multiple Mechanisms to Respond Abruptly to a Gradually Increasing Concentration of an Extracellular Signal

Some cell responses to extracellular signal molecules are smoothly graded according to the concentration of the signal molecule. In other cases, the relationship between signal and response can be discontinuous or all-or-none, with an abrupt switch from one type of outcome to another as the signal concentration increases beyond a certain value (**Figure 15–23**). Such effects are often the result of positive

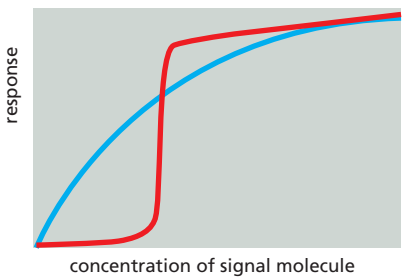


Figure 15–23 Smoothly graded versus switchlike signaling responses. Some cell responses increase gradually as the concentration of an extracellular signal molecule increases (*blue line*). In other cases, the responding cell is driven into an abruptly different state as the signal strength increases beyond a certain critical value *(red line)*.

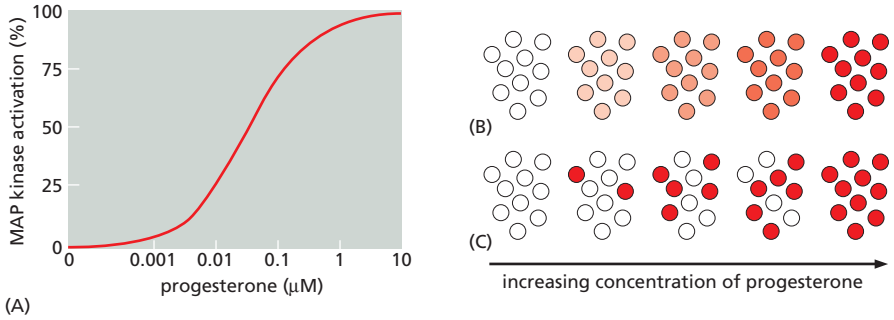


Figure 15–24 The importance of examining individual cells to detect all-or-none responses to increasing concentrations of an extracellular signal. In these experiments, immature frog eggs (oocytes) were stimulated with increasing concentrations of the hormone progesterone. The response was assessed by analyzing the activation of *MAP kinase* (discussed later), which is one of the protein kinases activated by phosphorylation in the response. The amount of phosphorylated (activated) MAP kinase in extracts of the oocytes was assessed biochemically. In (A), extracts of populations of stimulated oocytes were analyzed, and the activation of MAP kinase is seen to increase progressively with increasing progesterone concentration. There are two possible ways of explaining this result: (B) MAP kinase could have increased gradually in each individual cell with increasing progesterone concentration; (C) alternatively, individual cells could have responded in an all-or-none way, with the gradual increase in total MAP kinase activation reflecting the increasing number of cells responding with increasing progesterone concentration. When extracts of individual oocytes were analyzed, it was found that cells had either very low amounts or very high amounts, but not intermediate amounts, of the activated kinase, indicating that the response was essentially all-or-none at the level of individual cells, as diagrammed in (C). (Adapted from J.E. Ferrell and E.M. Machleder, *Science* 280:895–898, 1998. With permission from AAAS.)

feedback in the response system, as we discuss below. Both types of responses are common, and it is not always easy to distinguish between them. When one measures the effect of a signal on a whole population of cells, it may appear smoothly graded, even though the individual cells may be responding in an all-or-none manner but with variation from cell to cell in the signal concentration at which the switch occurs; **Figure 15–24** shows an example.

Moreover, smoothly graded responses can sometimes be very steeply dependent on the strength of the signal, giving the appearance of almost switchlike behavior. Cells use a variety of molecular mechanisms to acheive such effects. In one mechanism, more than one intracellular signaling molecule must bind to its downstream target protein to induce a response. As we discuss later, four molecules of the small intracellular mediator cyclic AMP, for example, must be bound simultaneously to each molecule of *cyclic-AMP-dependent protein kinase (PKA)* to activate the kinase. A similar sharpening of response is seen when the activation of an intracellular signaling protein requires phosphorylation at more than one site. Such *cooperative* responses become sharper as the number of required cooperating molecules or phosphate groups increases, and if the number is large enough, responses become almost all-or-none (**Figure 15–25**).

Responses are also sharpened when an intracellular signaling molecule activates one enzyme and, at the same time, inhibits another enzyme that catalyzes the opposite reaction. A well-studied example of this common type of regulation is the stimulation of glycogen breakdown in skeletal muscle cells induced by the hormone *adrenaline* (epinephrine). Adrenaline's binding to a G-protein-coupled cell-surface receptor increases the intracellular concentration of cyclic AMP, which both activates an enzyme that promotes glycogen breakdown and inhibits an enzyme that promotes glycogen synthesis.

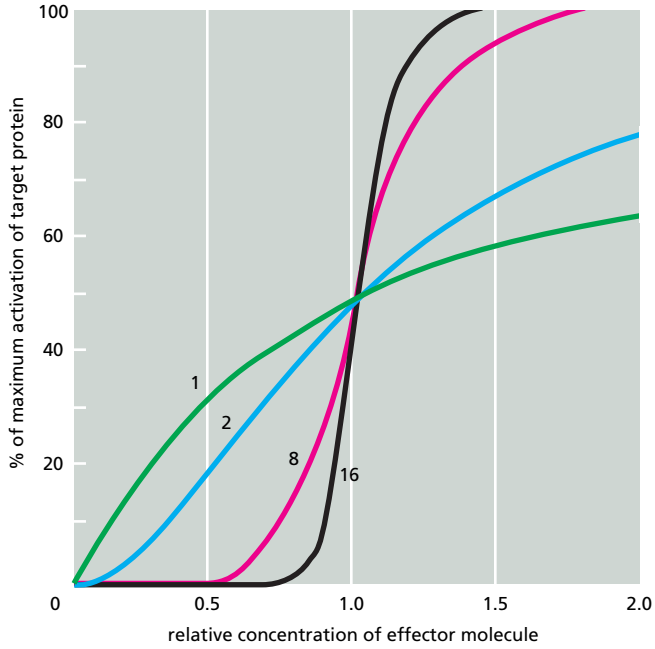


Figure 15–25 Activation curves for an allosteric protein as a function of effector molecule concentration. The curves show how the sharpness of the activation response increases with an increase in the number of allosteric effector molecules that must be bound simultaneously to activate the target protein. The curves shown are those expected, under certain conditions, if the activation requires the simultaneous binding of 1, 2, 8, or 16 effector molecules.

All of these mechanisms can produce very steep responses, but the responses are still smoothly graded according to the concentration of the extracellular signal molecule. To produce true all-or-none responses, one needs another mechanism, *positive feedback*, as we now discuss.

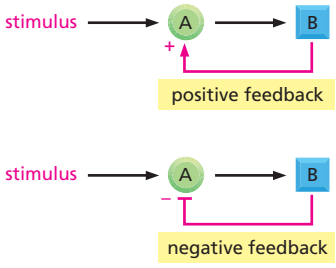


Figure 15–26 Positive and negative feedback. In these simple examples, a stimulus activates protein A, which, in turn, activates protein B. Protein B then acts back to either increase or decrease the activity of A.

Intracellular Signaling Networks Usually Make Use of Feedback Loops

Like intracellular metabolic pathways (discussed in Chapter 2), most intracellular signaling networks incorporate feedback loops, in which the output of a process acts back to regulate that same process. In *positive feedback,* the output stimulates its own production; in *negative feedback,* the output inhibits its own production (**Figure 15–26**). Feedback loops are of great general importance in biology, and they regulate many chemical and physical processes in cells. They operate over an enormous range of time scales in cells, from milliseconds (in the case of an action potential, for example—see Figure 11–29) to many hours (in the case of circadian oscillations, for example—see Figure 7–73). Those that regulate cell signaling can either operate exclusively within the target cell or involve the secretion of extracellular signals. Here, we focus on those feedback loops that operate entirely within the target cell; even the simplest of these loops can produce complex and interesting effects.

A **positive feedback loop** in a signaling pathway can transform the behavior of the responding cell. If the positive feedback is of only moderate strength, its effect will be simply to steepen and increase the response to the signal. But if the feedback is strong enough, it can produce a qualitatively different result: a runaway increase in the quantity of product when the signal increases above a critical value, leading to a new steady level of production that is sharply different from that obtained when the signal was only slightly weaker (**Figure 15–27**).

This type of all-or-none response goes hand in hand with a further property: once the responding system has switched to the high level of activation, this condition is self-sustaining and can persist even after the signal strength drops back below its critical value. In such a case, the system is said to be *bistable:* it can exist in either a "switched-off" or a "switched-on" state, and a transient stimulus can flip it from the one state to the other (**Figure 15–28**A and B). Cells use positive feedback loops of this sort to make stable all-or-none decisions, especially during development, when cells in different positions must choose between alternative developmental pathways in response to smoothly graded positional signals (morphogens), as discussed earlier. Through positive feedback, a transient extracellular signal can often induce long-term changes in cells and their progeny that can persist for the lifetime of the organism. The signals that trigger muscle cell specification, for example, turn on the transcription of a series of genes that encode muscle-specific gene regulatory proteins, which stimulate the transcription of their own genes, as well as genes encoding various other muscle cell proteins; in this way, the decision to become a muscle cell is made permanent (see Figure 7–75). This type of cell memory dependent on positive feedback is one of the basic ways in which a cell can undergo a lasting change of character without any alteration in its DNA sequence; and this altered state can be passed on to daughter cells. Such mechanisms of inheritance are called *epigenetic,* in contrast to genetic mechanisms involving mutations of the DNA, and they are discussed more fully in Chapter 7 (see Figure 7–86).

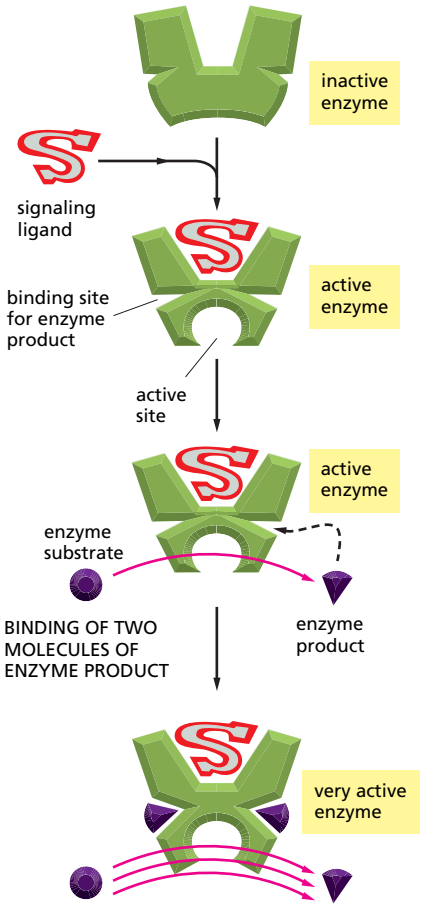


Figure 15–27 A positive feedback mechanism giving switch-like behavior. In this example, an intracellular signaling molecule (ligand) activates an enzyme located downstream in a signaling pathway. Two molecules of the product of the enzymatic reaction bind back to the same enzyme to activate it further. The consequence is a very low rate of synthesis of the product in the absence of the ligand. The rate increases slowly with the concentration of ligand until, at some threshold level of ligand, enough of the product has been synthesized to activate the enzyme in a self-accelerating, runaway fashion. The concentration of the product then rapidly increases to a much higher level.

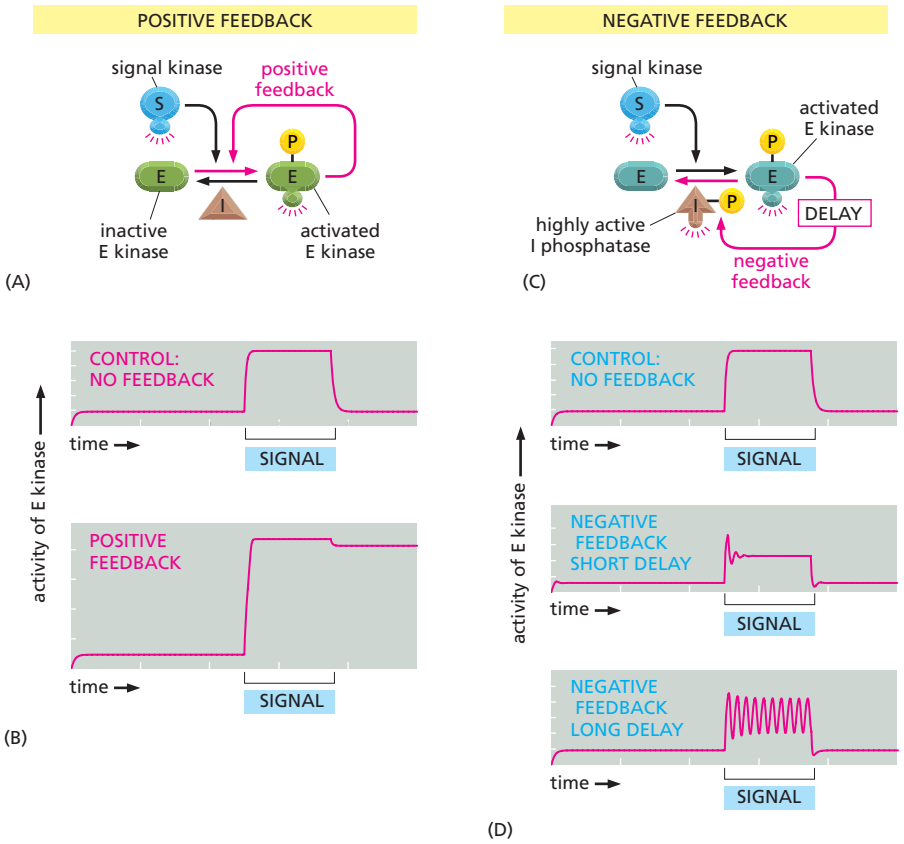


Figure 15–28 Some effects of simple feedback. The graphs show the computed effects of simple positive and negative feedback loops. In each case, the input signal is an activated protein kinase (S) that phosphorylates and thereby activates another protein kinase (E); a protein phosphatase (I) dephosphorylates and inactivates the activated E kinase. In the graphs, the *red* line indicates the activity of the E kinase over time; the underlying *blue* bar indicates the time for which the input signal (activated S kinase) is present. (A) Diagram of the positive feedback loop, in which the activated E kinase acts back to promote its own phosphorylation and activation; the basal activity of the I phosphatase dephosphorylates activated E at a steady low rate. (B) The top graph shows that, without feedback, the activity of the E kinase is simply proportional (with a short lag) to the level of stimulation by the S kinase. The bottom graph shows that, with the positive feedback loop, the system is bistable (i.e., it is capable of existing in either of two steady states): the transient stimulation by S kinase switches the system from an "off" state to an "on" state, which then persists after the stimulus has been removed. (C) Diagram of the negative feedback loop, in which the activated E kinase phosphorylates and activates the I phosphatase, thereby increasing the rate at which the phosphatase dephosphorylates and inactivates the phosphorylated E kinase. The top graph shows, again, the response in E kinase activity without feedback. The other graphs show the effects on E kinase activity of negative feedback operating after a short or long delay. With a short delay, the system shows a strong brief response when the signal is abruptly changed, and the feedback then drives the response back down to a lower level. With a long delay, the feedback produces sustained oscillations for as long as the stimulus is present.

Small changes in the details of the feedback can radically alter the way the system responds, even for this very simple example; the figure shows only a small sample of possible behaviors.

By contrast with positive feedback, a **negative feedback loop** counteracts the effect of a stimulus and thereby abbreviates and limits the level of the response, making the system less sensitive to perturbations. As with positive feedback, however, qualitatively different phenomena can be obtained when the feedback operates more powerfully. A delayed negative feedback with a long enough delay can produce responses that oscillate. The oscillations may persist for as long as the stimulus is present (Figure 15–28C) or they may even be generated spontaneously, without need of an external signal to drive them (see Figure 22–82). Later in this chapter, we shall encounter a number of examples of such oscillatory behavior in the intracellular responses to extracellular signals; all of them depend on negative feedback.

If negative feedback operates with a short delay, the system behaves like a change detector. It gives a strong response to a stimulus, but the response rapidly decays even while the stimulus persists; if the stimulus is suddenly increased, however, the system responds strongly again, but, again, the response rapidly decays. This is the phenomenon of *adaptation,* which we now discuss.

Cells Can Adjust Their Sensitivity to a Signal

In responding to many types of stimuli, cells and organisms are able to detect the same percentage of change in a signal over a very wide range of stimulus strengths. The target cells accomplish this through a reversible process of **adaptation**, or **desensitization**, whereby a prolonged exposure to a stimulus decreases the cells' response to that level of stimulus. In chemical signaling, adaptation enables cells to respond to *changes* in the concentration of an extracellular signal molecule (rather than to the absolute concentration of the signal) over a very wide range of signal concentrations. The underlying mechanism is negative feedback that operates with a short delay: a strong response modifies the signaling machinery involved, such that the machinery resets itself to become less responsive to the same level of signal (see Figure 15–28D, middle graph). Owing to the delay, however, a sudden increase in the signal is able to stimulate the cell again for a short period before the negative feedback has time to kick in.

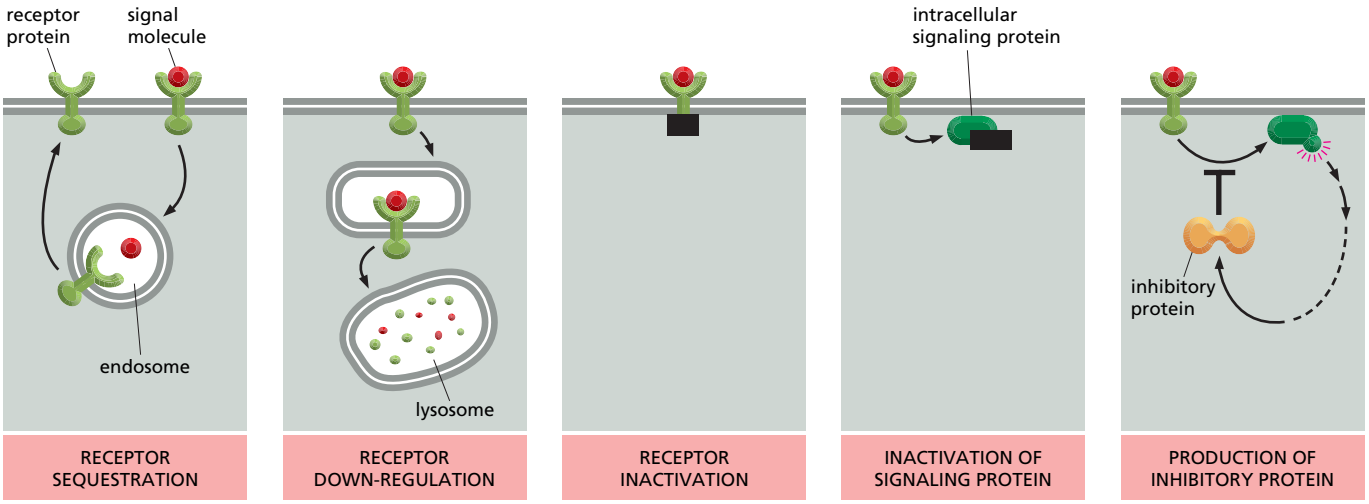


Figure 15–29 Some ways in which target cells can become desensitized (adapted) to an extracellular signal molecule. The mechanisms shown here that operate at the level of the receptor often involve phosphorylation or ubiquitylation of the receptor proteins. In bacterial chemotaxis, however, which we discuss later, adaptation depends on methylation of the receptor proteins.

Desensitization to a signal molecule can occur in various ways. It can result from an inactivation of the receptors themselves. The binding of signal molecules to cell-surface receptors, for example, may induce the endocytosis and temporary sequestration of the receptors in endosomes. In some cases, such signal-induced receptor endocytosis leads to the destruction of the receptors in lysosomes, a process referred to as *receptor down-regulation* (in other cases, however, activated receptors continue to signal after they have been endocytosed, as we discuss later). Receptors can also become desensitized (inactivated) on the cell surface—for example, by becoming phosphorylated or methylated—with a short delay following their activation. Desensitization can also occur at sites downstream of the receptors, either by a change in intracellular signaling proteins involved in transducing the extracellular signal or by the production of an inhibitor protein that blocks the signal transduction process. These various desensitization mechanisms are compared in **Figure 15–29**.

Having discussed some of the general principles of cell signaling, we next turn to the G-protein-coupled receptors, which are by far the largest class of cell-surface receptors, mediating responses to the most diverse types of extracellular signals.

Summary

Each cell in a multicellular animal has been programmed during development to respond to a specific set of extracellular signal molecules produced by other cells. The signal molecules act in various combinations to regulate the behavior of the cell. Most signal molecules act as local mediators, which are secreted and then rapidly taken up, destroyed, or immobilized, so that they act only on neighboring cells. Other signal molecules remain bound to the surface of the signaling cell and mediate contact-dependent signaling. There are also two distinct types of long-distance signaling. In endocrine signaling, hormones secreted by endocrine cells are carried in the blood to target cells throughout the body. In synaptic signaling, neurotransmitters secreted by nerve cell axons act locally on the postsynaptic cells that the axons contact.

Cell signaling requires not only extracellular signal molecules but also a complementary set of receptor proteins expressed by the target cells that specifically bind the signal molecules. Some small hydrophobic signal molecules, including steroid and thyroid hormones, diffuse across the plasma membrane of the target cell and activate intracellular receptor proteins that directly regulate the transcription of specific genes. The dissolved gases nitric oxide and carbon monoxide act as local mediators by diffusing across the plasma membrane of the target cell and activating an intracellular protein such as the enzyme guanylyl cyclase, which produces cyclic GMP in the target cell. But most extracellular signal molecules are hydrophilic and cannot pass through the plasma membrane. These activate cell-surface receptor proteins, which

act as signal transducers, converting the extracellular signal into intracellular ones that alter the behavior of the target cell.

The three largest families of cell-surface receptors transduce extracellular signals in different ways. Ion-channel-coupled receptors are transmitter-gated ion channels that open or close briefly in response to the binding of a neurotransmitter. G-protein-coupled receptors indirectly activate or inactivate plasma-membrane-bound enzymes or ion channels via trimeric GTP-binding proteins (G proteins). Enzyme-coupled receptors either act directly as enzymes or are associated with enzymes; these enzymes are usually protein kinases that phosphorylate the receptors and specific signaling proteins in the target cell.

Once activated, G-protein-coupled receptors and enzyme-coupled receptors relay the signal into the cell interior by activating chains of intracellular signaling proteins. Some of these signaling proteins transduce, amplify, or spread the signal as they relay it, while others integrate signals from different signaling pathways. Some function as switches that are transiently activated by phosphorylation or GTP binding. Large signaling complexes form by means of modular interaction domains in the signaling proteins, which allow the proteins to form functional signaling networks.

Target cells use various intracellular mechanisms, including feedback loops, to adjust the ways in which they respond to extracellular signals. Positive feedback loops can help cells to respond in an all-or-none fashion to a gradually increasing concentration of an extracellular signal or to convert a short-lasting signal into a long-lasting, or even irreversible, response. Delayed negative feedback allows cells to desensitize to a signal molecule, which enables them to respond to small changes in the concentration of the signal molecule over a large concentration range.

SIGNALING THROUGH G-PROTEIN-COUPLED CELL-SURFACE RECEPTORS (GPCRs) AND SMALL INTRACELLULAR MEDIATORS

All eucaryotes use **G-protein-coupled receptors (GPCRs)**. These form the largest family of cell-surface receptors, and they mediate most responses to signals from the external world, as well as signals from other cells, including hormones, neurotransmitters, and local mediators. Our senses of sight, smell, and taste (with the possible exception of sour taste) depend on them. There are more than 700 GPCRs in humans, and in mice there are about 1000 concerned with the sense of smell alone. The signal molecules that act on GPCRs are as varied in structure as they are in function and include proteins and small peptides, as well as derivatives of amino acids and fatty acids, not to mention photons of light and all the molecules that we can smell or taste. The same signal molecule can activate many different GPCR family members; for example, adrenaline activates at least 9 distinct GPCRs, acetylcholine another 5, and the neurotransmitter serotonin at least 14. The different receptors for the same signal are usually expressed in different cell types and elicit different responses.

Despite the chemical and functional diversity of the signal molecules that activate them, all GPCRs have a similar structure. They consist of a single polypeptide chain that threads back and forth across the lipid bilayer seven times (**Figure 15–30**). In addition to their characteristic orientation in the plasma membrane, they all use G proteins to relay the signal into the cell interior.

The GPCR superfamily includes *rhodopsin,* the light-activated protein in the vertebrate eye, as well as the large number of olfactory receptors in the vertebrate nose. Other family members are found in unicellular organisms: the receptors in yeasts that recognize secreted mating factors are an example. It is likely that the GPCRs that mediate cell–cell signaling in multicellular organisms evolved from the sensory receptors in their unicellular eucaryotic ancestors.

It is remarkable that about half of all known drugs work through GPCRs or the signaling pathways GPCRs activate. Of the many hundreds of genes in the human genome that encode GPCRs, about 150 encode orphan receptors, for which the ligand is unknown. Many of them are likely targets for new drugs that remain to be discovered.

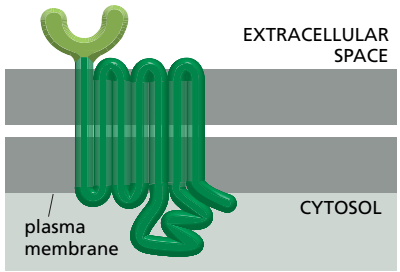


Figure 15–30 A G-protein-coupled receptor (GPCR). GPCRs that bind protein ligands have a large extracellular domain formed by the part of the polypeptide chain shown in *light green.* This domain, together with some of the transmembrane segments, binds the protein ligand. Receptors for small ligands such as adrenaline have small extracellular domains, and the ligand usually binds deep within the plane of the membrane to a site that is formed by amino acids from several transmembrane segments.

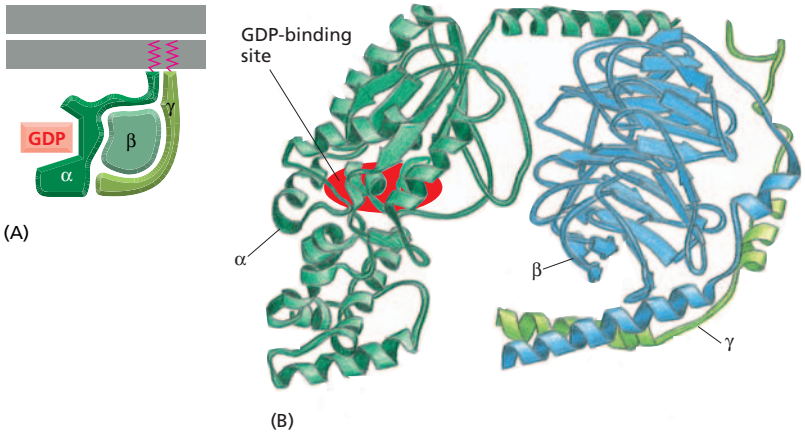


Figure 15–31 The structure of an inactive G protein. (A) Note that both the α and the γ subunits have covalently attached lipid molecules *(red)* that help bind them to the plasma membrane, and the α subunit has GDP bound. (B) The three-dimensional structure of an inactive G protein, based on transducin, the G protein that operates in visual transduction (discussed later). The α subunit contains the GTPase domain and binds to one side of the β subunit, which locks the GTPase domain in an inactive conformation that binds GDP. The γ subunit binds to the opposite side of the β subunit, and the β and γ subunits together form a single functional unit. (B, based on D.G. Lombright et al., *Nature* 379:311–319, 1996. With permission from Macmillan Publishers Ltd.)

Trimeric G Proteins Relay Signals from GPCRs

When an extracellular signal molecule binds to a GPCR, the receptor undergoes a conformational change that enables it to activate a **trimeric GTP-binding protein (G protein)**. The G protein is attached to the cytoplasmic face of the plasma membrane, where it functionally couples the receptor to either enzymes or ion channels in this membrane. In some cases, the G protein is physically associated with the receptor before the receptor is activated, whereas in others it binds only after receptor activation. There are various types of G proteins, each specific for a particular set of GPCRs and for a particular set of target proteins in the plasma membrane. They all have a similar structure, however, and operate similarly.

G proteins are composed of three protein subunits—α, β, and γ. In the unstimulated state, the α subunit has GDP bound and the G protein is inactive (**Figure 15–31**). When a GPCR is activated, it acts like a guanine nucleotide exchange factor (GEF) and induces the α subunit to release its bound GDP, allowing GTP to bind in its place. This exchange causes a large conformational change in the G protein, which activates it. It was originally thought that the activation always causes the trimer to dissociate into two activated components—an *α subunit* and a *βγ complex*. However, there is now evidence that, in some cases at least, the conformational change exposes previously buried surfaces between the α subunit and the βγ complex, so that the α subunit and βγ complex can each now interact with their targets without requiring the subunits to dissociate (**Figure 15–32**). These targets are either enzymes or ion channels in the plasma membrane that relay the signal onward.

The α subunit is a GTPase, and once it hydrolyzes its bound GTP to GDP it becomes inactive. The time for which the G protein remains active depends on how quickly the α subunit hydrolyzes its bound GTP. This time is usually short because the GTPase activity is greatly enhanced by the binding of the α subunit to a second protein, which can be either the target protein or a specific **regulator of G protein signaling** (**RGS**). RGS proteins act as α-subunit-specific GTPase-activating proteins (GAPs) (see Figure 15–19), and they help shut off G-protein-mediated responses in all eucaryotes. There are about 25 RGS proteins encoded in the human genome, each of which interacts with a particular set of G proteins.

GPCRs activate various intracellular signaling pathways, including some that are also activated by enzyme-coupled receptors. In this section, however, we focus on those GPCR-activated pathways that use small intracellular mediators.

Some G Proteins Regulate the Production of Cyclic AMP

Cyclic AMP (cAMP) acts as a small intracellular mediator in all procaryotic and animal cells that have been studied. Its normal concentration in the cytosol is about 10^{-7} M, but an extracellular signal can increase this concentration more than twentyfold in seconds (**Figure 15–33**). As explained earlier (see Figure

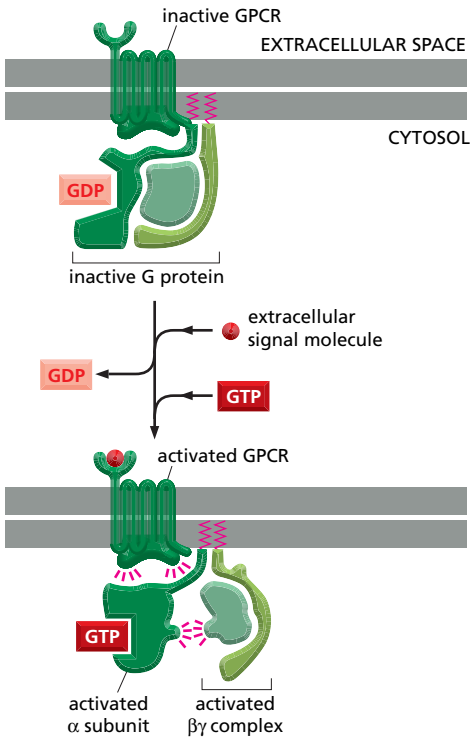


Figure 15–32 Activation of a G protein by an activated GPCR. Binding of an extracellular signal to a GPCR changes the conformation of the receptor, which in turn alters the conformation of the G protein. The alteration of the α subunit of the G protein allows it to exchange its GDP for GTP, activating both the α subunit and the βγ complex, both of which can regulate the activity of target proteins in the plasma membrane. The receptor stays active while the external signal molecule is bound to it, and it can therefore catalyze the activation of many molecules of G protein, which dissociate from the receptor once activated (not shown). In some cases, the α subunit and the βγ complex dissociate from each other when the G protein is activated.

15–11), such a rapid response requires balancing a rapid synthesis of the molecule with its rapid breakdown or removal. Cyclic AMP is synthesized from ATP by a plasma-membrane-bound enzyme **adenylyl cyclase**, and it is rapidly and continuously destroyed by **cyclic AMP phosphodiesterases** that hydrolyze cyclic AMP to adenosine 5′-monophosphate (5′-AMP) (**Figure 15–34**).

Many extracellular signal molecules work by increasing cyclic AMP concentration, and they do so by increasing the activity of adenylyl cyclase against a steady background of phosphodiesterase activity. Adenylyl cyclase is a large multipass transmembrane protein with its catalytic domain on the cytosolic side of the plasma membrane. There are at least eight isoforms in mammals, most of which are regulated by both G proteins and Ca^{2+}. GPCRs that act by increasing cyclic AMP are coupled to a **stimulatory G protein (G_s)**, which activates adenylyl cyclase and thereby increases cyclic AMP concentration. Another G protein, called **inhibitory G protein (G_i)**, inhibits adenylyl cyclase, but it acts mainly by directly regulating ion channels (as we discuss later).

Both G_s and G_i are targets for some medically important bacterial toxins. *Cholera toxin*, which is produced by the bacterium that causes cholera, is an enzyme that catalyzes the transfer of ADP ribose from intracellular NAD^+ to the α subunit of G_s. This ADP ribosylation alters the α subunit so that it can no longer hydrolyze its bound GTP, causing it to remain in an active state that stimulates adenylyl cyclase indefinitely. The resulting prolonged elevation in cyclic AMP concentration within intestinal epithelial cells causes a large efflux of Cl^- and water into the gut, thereby causing the severe diarrhea that characterizes cholera. *Pertussis toxin*, which is made by the bacterium that causes pertussis

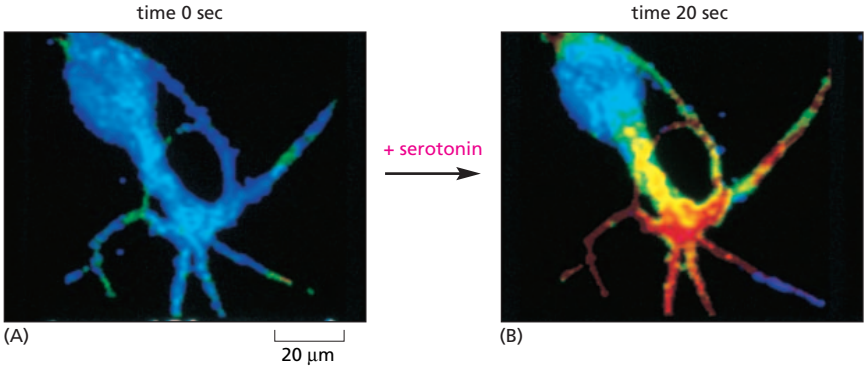


Figure 15–33 An increase in cyclic AMP in response to an extracellular signal. This nerve cell in culture is responding to the neurotransmitter serotonin, which acts through a GPCR to cause a rapid rise in the intracellular concentration of cyclic AMP. To monitor the cyclic AMP level, the cell has been loaded with a fluorescent protein that changes its fluorescence when it binds cyclic AMP. *Blue* indicates a low level of cyclic AMP, *yellow* an intermediate level, and *red* a high level. (A) In the resting cell, the cyclic AMP level is about 5×10^{-8} M. (B) Twenty seconds after the addition of serotonin to the culture medium, the intracellular level of cyclic AMP has increased to more than 10^{-6} M in the relevant parts of the cell, an increase of more than twentyfold. (From Brian J. Bacskai et al., *Science* 260:222–226, 1993. With permission from AAAS.)

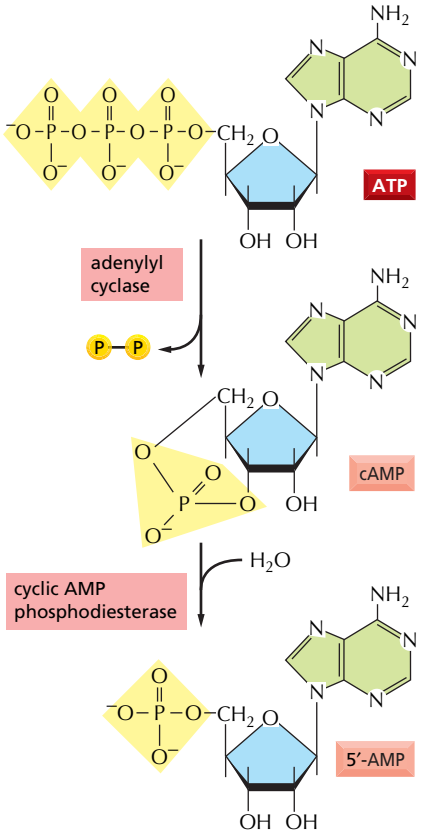


Figure 15–34 The synthesis and degradation of cyclic AMP. In a reaction catalyzed by the enzyme adenylyl cyclase, cyclic AMP (cAMP) is synthesized from ATP through a cyclization reaction that removes two phosphate groups as pyrophosphate (Ⓟ – Ⓟ); a pyrophosphatase drives this synthesis by hydrolyzing the released pyrophosphate to phosphate (not shown). Cyclic AMP is short-lived (unstable) in the cell because it is hydrolyzed by specific phosphodiesterases to form 5′-AMP, as indicated.

(whooping cough), catalyzes the ADP ribosylation of the α subunit of G_i, preventing the protein from interacting with receptors; as a result, the G protein retains its bound GDP and is unable to regulate its target proteins. These two toxins are widely used in experiments to determine whether a cell's GPCR-dependent response to a signal is mediated by G_s or by G_i.

Some of the responses mediated by a G_s-stimulated increase in cyclic AMP concentration are listed in **Table 15–1**. As the table shows, different cell types respond differently to an increase in cyclic AMP concentration, and one cell type often responds in the same way to such an increase, regardless of the extracellular signal that causes it. At least four hormones activate adenylyl cyclase in fat cells, for example, and all of them stimulate the breakdown of triglyceride (the storage form of fat) to fatty acids (see Table 15–1).

Individuals who are genetically deficient in a particular G_s α subunit show decreased responses to certain hormones. As a consequence, these people display metabolic abnormalities, have abnormal bone development, and are mentally retarded.

Table 15–1 Some Hormone-induced Cell Responses Mediated by Cyclic AMP

TARGET TISSUE	HORMONE	MAJOR RESPONSE
Thyroid gland	thyroid-stimulating hormone (TSH)	thyroid hormone synthesis and secretion
Adrenal cortex	adrenocorticotrophic hormone (ACTH)	cortisol secretion
Ovary	luteinizing hormone (LH)	progesterone secretion
Muscle	adrenaline	glycogen breakdown
Bone	parathormone	bone resorption
Heart	adrenaline	increase in heart rate and force of contraction
Liver	glucagon	glycogen breakdown
Kidney	vasopressin	water resorption
Fat	adrenaline, ACTH, glucagon, TSH	triglyceride breakdown

Cyclic-AMP-Dependent Protein Kinase (PKA) Mediates Most of the Effects of Cyclic AMP

In most animal cells, cyclic AMP exerts its effects mainly by activating **cyclic-AMP-dependent protein kinase (PKA)**. This kinase phosphorylates specific serines or threonines on selected target proteins, including intracellular signaling proteins and effector proteins, thereby regulating their activity. The target proteins differ from one cell type to another, which explains why the effects of cyclic AMP vary so markedly depending on the cell type (see Table 15–1).

In the inactive state, PKA consists of a complex of two catalytic subunits and two regulatory subunits. The binding of cyclic AMP to the regulatory subunits alters their conformation, causing them to dissociate from the complex. The released catalytic subunits are thereby activated to phosphorylate specific target proteins (**Figure 15–35**). The regulatory subunits of PKA (also called A-kinase) are important for localizing the kinase inside the cell: special *A-kinase anchoring proteins (AKAPs)* bind both to the regulatory subunits and to a component of the cytoskeleton or a membrane of an organelle, thereby tethering the enzyme complex to a particular subcellular compartment. Some AKAPs also bind other signaling proteins, forming a complex that functions as a signaling module. An AKAP located around the nucleus of heart muscle cells, for example, binds both PKA and a phosphodiesterase that hydrolyzes cyclic AMP. In unstimulated cells, the phosphodiesterase keeps the local cyclic AMP concentration low, so that the bound PKA is inactive; in stimulated cells, cyclic AMP concentration rapidly rises, overwhelming the phosphodiesterase and activating the PKA. Among the target proteins that PKA phosphorylates and activates in these cells is the adjacent phosphodiesterase, which rapidly lowers the cyclic AMP concentration again. This arrangement converts what might otherwise be a weak and prolonged PKA response into a strong, brief, local pulse of PKA activity.

Whereas some responses mediated by cyclic AMP occur within seconds and do not depend on changes in gene transcription (see Figure 15–33), others do depend on changes in the transcription of specific genes and take hours to develop fully. In cells that secrete the peptide hormone *somatostatin,* for example, cyclic AMP activates the gene that encodes this hormone. The regulatory region of the somatostatin gene contains a short DNA sequence, called the *cyclic AMP response element (CRE),* which is also found in the regulatory region of many other genes activated by cyclic AMP. A specific gene regulatory protein called **CRE-binding (CREB) protein** recognizes this sequence. When PKA is activated by cAMP, it phosphorylates CREB on a single serine; phosphorylated CREB then recruits a transcriptional coactivator called *CREB-binding protein (CBP),* which stimulates the transcription of the target genes (**Figure 15–36**). Thus, CREB can transform a short cyclic AMP signal into a long-term change in a cell, a process that, in the brain, is thought to play an important part in some forms of learning and memory.

PKA does not mediate all the effects of cyclic AMP in cells. As we discuss later, in olfactory neurons (responsible for the sense of smell), cyclic AMP also directly activates special ion channels in the plasma membrane. Moreover, in some other

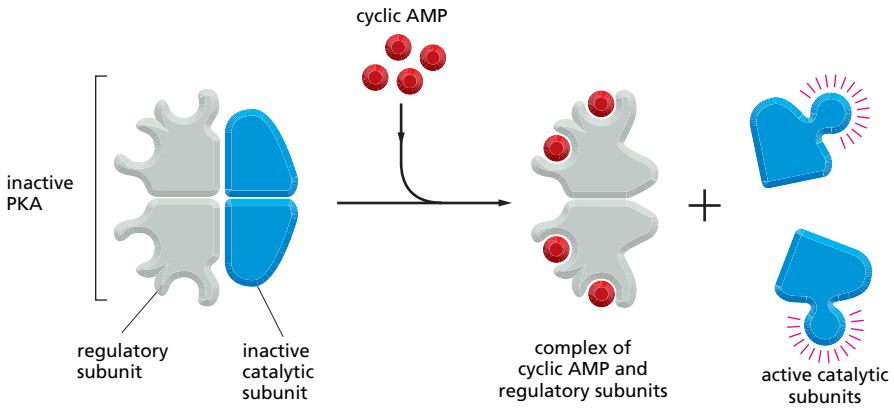


Figure 15–35 The activation of cyclic-AMP-dependent protein kinase (PKA). The binding of cyclic AMP to the regulatory subunits of the PKA tetramer induces a conformational change, causing these subunits to dissociate from the catalytic subunits, thereby activating the kinase activity of the catalytic subunits. The release of the catalytic subunits requires the binding of more than two cyclic AMP molecules to the regulatory subunits in the tetramer. This requirement greatly sharpens the response of the kinase to changes in cyclic AMP concentration, as discussed earlier (see Figure 15–25). Mammalian cells have at least two types of PKAs: type I is mainly in the cytosol, whereas type II is bound via its regulatory subunits and special anchoring proteins to the plasma membrane, nuclear membrane, mitochondrial outer membrane, and microtubules. In both types, once the catalytic subunits are freed and active, they can migrate into the nucleus (where they can phosphorylate gene regulatory proteins), while the regulatory subunits remain in the cytoplasm. The three-dimensional structure of the protein kinase domain of the PKA catalytic subunit is shown in Figure 3–65.

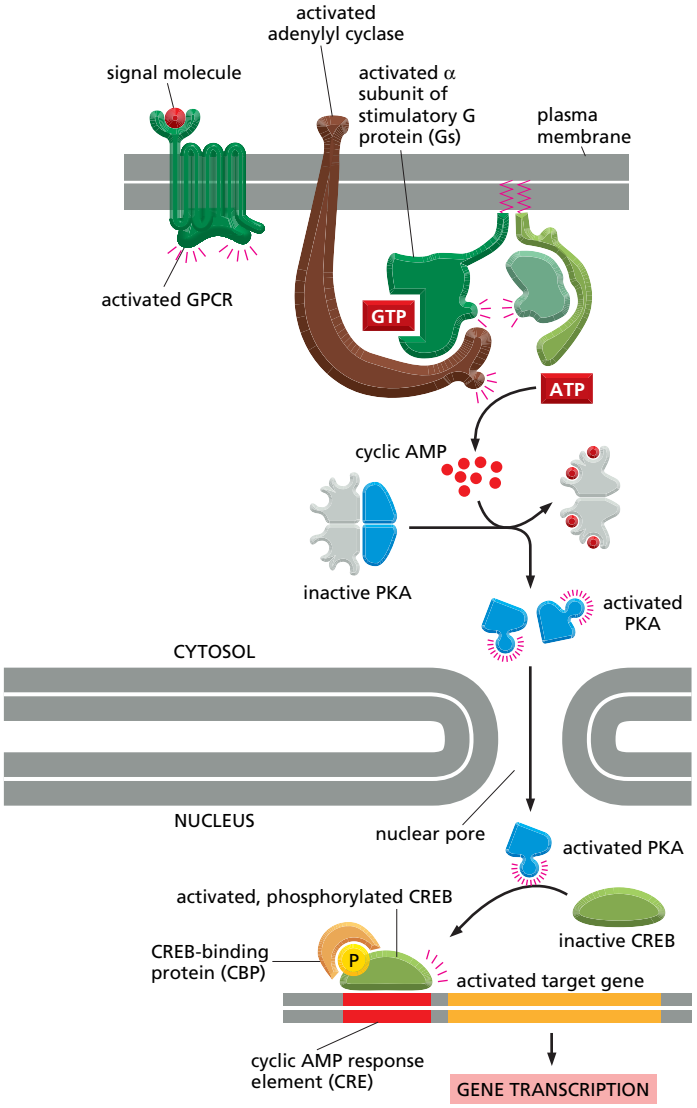


Figure 15–36 How a rise in intracellular cyclic AMP concentration can alter gene transcription. <AGAT> The binding of an extracellular signal molecule to its GPCR activates adenylyl cyclase via G_s and thereby increases cyclic AMP concentration in the cytosol. The rise in cyclic AMP concentration activates PKA, and the released catalytic subunits of PKA can then enter the nucleus, where they phosphorylate the gene regulatory protein CREB. Once phosphorylated, CREB recruits the coactivator CBP, which stimulates gene transcription. In some cases, at least, the inactive CREB protein is bound to the cyclic AMP response element (CRE) in DNA before it is phosphorylated (not shown).

This signaling pathway controls many processes in cells, ranging from hormone synthesis in endocrine cells to the production of proteins required for the induction of long-term memory in the brain. We will see later that CREB can also be activated by some other signaling pathways, independent of cAMP.

cells, it directly activates a guanine nucleotide exchange factor (GEF) that, in turn, activates a monomeric GTPase called *Rap1*, which often leads to increased cell adhesion through the activation of cell-surface *integrins* (discussed in Chapter 19).

Having discussed how trimeric G proteins link activated GPCRs to adenylyl cyclase, we now consider how they couple activated GPCRs to another crucial enzyme, *phospholipase C*. The activation of this enzyme increases the concentration of several small intracellular mediators, including Ca^{2+}, which help relay the signal onward. Ca^{2+} is even more widely used than cyclic AMP as a small intracellular mediator.

Some G Proteins Activate An Inositol Phospholipid Signaling Pathway by Activating Phospholipase C-β

Many GPCRs exert their effects mainly via G proteins that activate the plasma-membrane-bound enzyme **phospholipase C-β (PLCβ)**. **Table 15–2** lists several examples of responses activated in this way. The phospholipase acts on a phosphorylated inositol phospholipid (a *phosphoinositide*) called **phosphatidylinositol 4,5-bisphosphate [PI(4,5)P_2, or PIP_2]**, which is present in small amounts in the inner half of the plasma membrane lipid bilayer (**Figure 15–37**). Receptors that activate this **inositol phospholipid signaling pathway** mainly do so via a G protein called $\mathbf{G_q}$, which activates phospholipase C-β in much the same way that

Table 15–2 Some Cell Responses in Which GPCRs Activate PLCβ

TARGET TISSUE	SIGNAL MOLECULE	MAJOR RESPONSE
Liver	vasopressin	glycogen breakdown
Pancreas	acetylcholine	amylase secretion
Smooth muscle	acetylcholine	muscle contraction
Blood platelets	thrombin	platelet aggregation

G_s activates adenylyl cyclase. The activated phospholipase then cleaves the PIP_2 to generate two products: *inositol 1,4,5-trisphosphate (IP_3)* and *diacylglycerol* (**Figure 15–38**). At this step, the signaling pathway splits into two branches.

Inositol 1,4,5-trisphosphate (IP_3) is a water-soluble molecule that acts as a small intracellular mediator. It leaves the plasma membrane and diffuses rapidly through the cytosol. When it reaches the endoplasmic reticulum (ER), it binds to and opens **IP_3-gated Ca^{2+}-release channels** (also called **IP_3 receptors**) in the ER membrane. Ca^{2+} stored in the ER is released through the open channels, quickly raising the concentration of Ca^{2+} in the cytosol (**Figure 15–39**). Once the ER Ca^{2+} stores have been depleted, they are refilled by the activation of *store-operated Ca^{2+} channels* in the plasma membrane and a Ca^{2+} sensor protein in the ER membrane, in regions where the two membranes are closely apposed.

We discuss later how the increase in cytosolic Ca^{2+} propagates the signal by influencing the activity of Ca^{2+}-sensitive intracellular proteins. Several mechanisms operate to terminate the initial Ca^{2+} response: (1) IP_3 is rapidly dephosphorylated by specific lipid phosphatases to form IP_2; (2) IP_3 is phosphorylated by specific lipid kinases to form IP_4 (which may function as another small intracellular mediator); and (3) Ca^{2+} that enters the cytosol is rapidly pumped out, mainly to the exterior of the cell (see Figure 15–41).

At the same time that the IP_3 produced by the hydrolysis of PIP_2 is increasing the concentration of Ca^{2+} in the cytosol, the other cleavage product of the PIP_2, **diacylglycerol**, is exerting different effects. It also acts as a small intracellular mediator, but it remains embedded in the plasma membrane, where it has several potential signaling roles. It can be further cleaved to release arachidonic acid, which can either act as a signal in its own right or be used in the synthesis of other small lipid signal molecules called *eicosanoids*. Most vertebrate cell

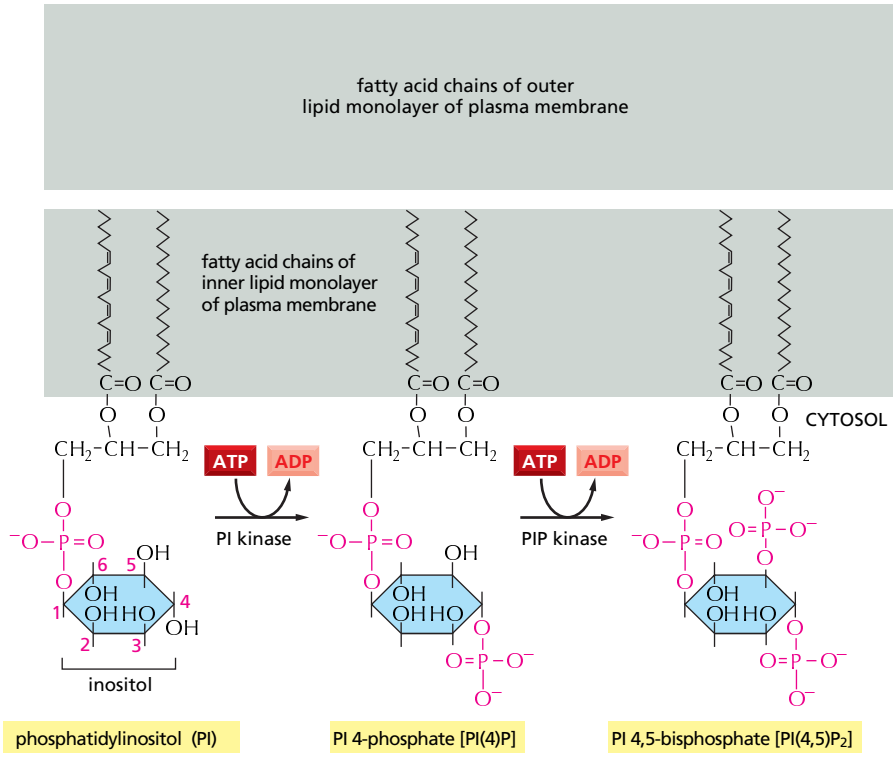


Figure 15–37 The synthesis of PI(4,5)P_2. The phosphoinositides PI(4)P and PI(4,5)P_2 are produced by the phosphorylation of phosphatidylinositol (PI) and PI(4)P, respectively. Although all three inositol phospholipids may be broken down in a signaling response, it is the breakdown of PI(4,5)P_2 that predominates and is most critical because it generates two intracellular mediators, as shown in Figures 15–38 and 15–39. Nevertheless, PI(4,5)P_2 is the least abundant, constituting less than 10% of the total inositol phospholipids and about 1% of the lipids in the plasma membrane. The conventional numbering of the carbon atoms in the inositol ring is shown in *red numbers* on the PI molecule.

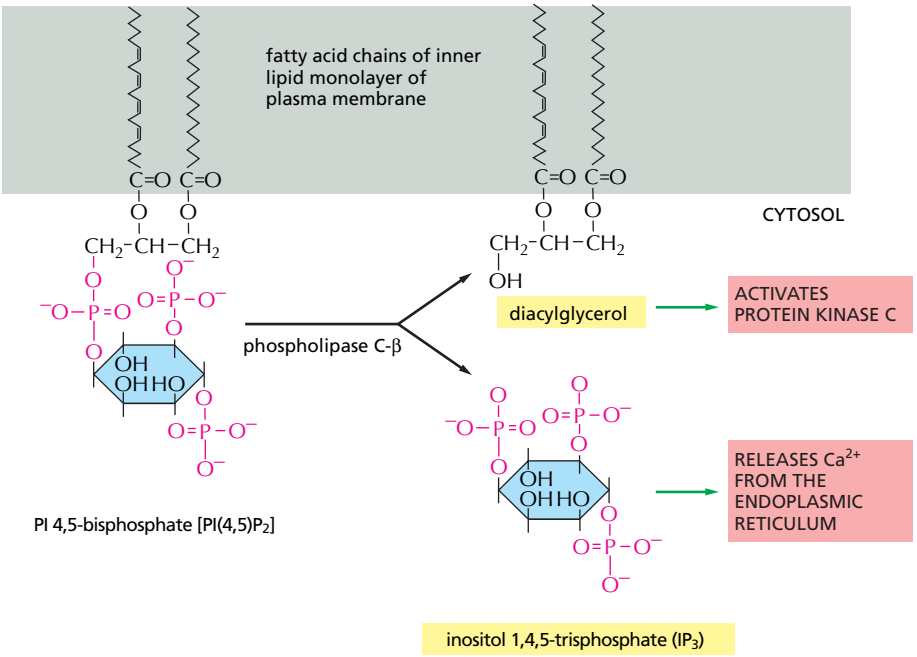


Figure 15–38 The hydrolysis of PI(4,5)P_2 by phospholipase C-β. Two small intracellular mediators are produced directly from the hydrolysis of the PIP_2: inositol 1,4,5-trisphosphate (IP_3), which diffuses through the cytosol and releases Ca^{2+} from the endoplasmic reticulum, and diacylglycerol, which remains in the membrane and helps to activate the enzyme protein kinase C (PKC; see Figure 15–39). There are several classes of phospholipase C; these include the β class, which is activated by GPCRs; as we see later, the γ class is activated by a class of enzyme-coupled receptors called receptor tyrosine kinases (RTKs).

types make eicosanoids, including *prostaglandins,* which have many biological activities. They participate in pain and inflammatory responses, for example, and most anti-inflammatory drugs (such as aspirin, ibuprofen, and cortisone) act—in part, at least—by inhibiting their synthesis.

A second function of diacylglycerol is to activate a crucial serine/threonine protein kinase called **protein kinase C (PKC)**, so named because it is Ca^{2+}-dependent. The initial rise in cytosolic Ca^{2+} induced by IP_3 alters the PKC so that it translocates from the cytosol to the cytoplasmic face of the plasma membrane. There it is activated by the combination of Ca^{2+}, diacylglycerol, and the negatively charged membrane phospholipid phosphatidylserine (see Figure 15–39). Once activated, PKC phosphorylates target proteins that vary depending on the cell type. The principles are the same as discussed earlier for PKA, although most of the target proteins are different.

There are various classes of PKCs, only some of which (called *conventional PKCs*) are activated by Ca^{2+} and diacylglycerol; the others are called *atypical PKCs.* Different PKCs phosphorylate different substrates mainly because different anchoring or scaffold proteins tether them to different compartments in the cell.

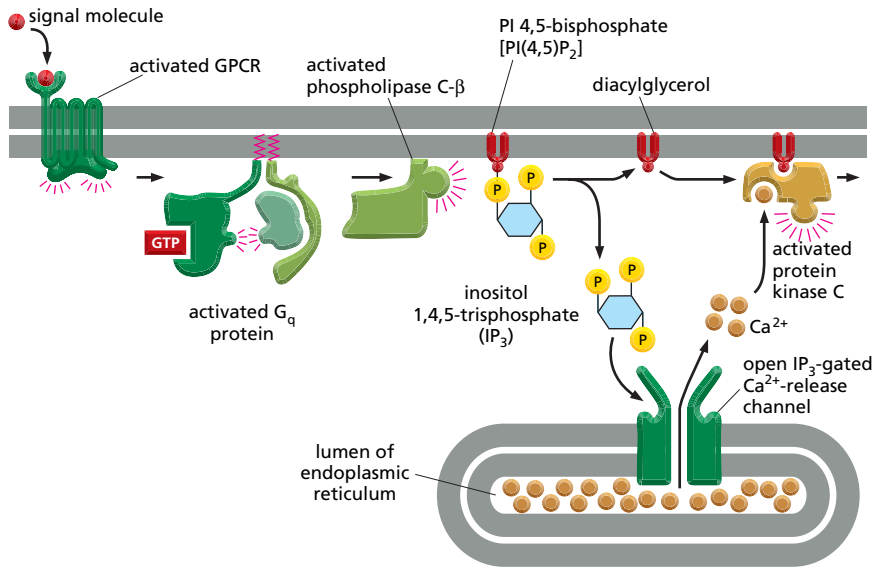


Figure 15–39 How GPCRs increase cytosolic Ca^{2+} and activate PKC. The activated GPCR stimulates the plasma-membrane-bound phospholipase PLCβ via a G protein. Depending on the isoform of the PLCβ, it may be activated by the α subunit of G_q as shown, by the βγ subunits of another G protein, or by both. Two small intracellular messenger molecules are produced when PI(4,5)P_2 is hydrolyzed by activated PLCβ. Inositol 1,4,5-trisphosphate (IP_3) diffuses through the cytosol and releases Ca^{2+} from the ER by binding to and opening IP_3-gated Ca^{2+}-release channels (IP_3 receptors) in the ER membrane. The large electrochemical gradient for Ca^{2+} across this membrane causes Ca^{2+} to escape into the cytosol when the release channels are open. Diacylglycerol remains in the plasma membrane and, together with phosphatidylserine (not shown) and Ca^{2+}, helps to activate protein kinase C (PKC), which is recruited from the cytosol to the cytosolic face of the plasma membrane. Of the 10 or more distinct isoforms of PKC in humans, at least 4 are activated by diacylglycerol.

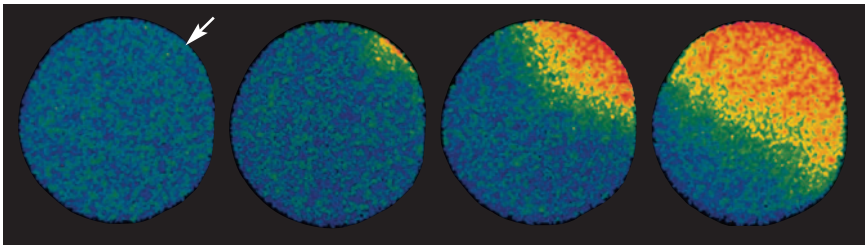


Figure 15–40 The fertilization of an egg by a sperm triggers an increase in cytosolic Ca^{2+}. <AGGA> This starfish egg was injected with a Ca^{2+}-sensitive fluorescent dye before it was fertilized. A wave of cytosolic Ca^{2+} *(red),* released from the ER, sweeps across the egg from the site of sperm entry *(arrow)*. This Ca^{2+} wave changes the egg cell surface, preventing the entry of other sperm, and it also initiates embryonic development (discussed in Chapter 21). The initial increase in Ca^{2+} is thought to be caused by a sperm-specific form of PLC (PLCζ) that the sperm brings into the egg cytoplasm when it fuses with the egg; the PLCζ cleaves $PI(4,5)P_2$ to produce IP_3, which releases Ca^{2+} from the egg ER. (Courtesy of Stephen A. Stricker.)

Ca^{2+} Functions as a Ubiquitous Intracellular Mediator

Many extracellular signals trigger an increase in cytosolic Ca^{2+} concentration, not just those that work via G proteins. In egg cells, for example, a sudden rise in cytosolic Ca^{2+} concentration upon fertilization by a sperm triggers a Ca^{2+} wave that initiates embryonic development (**Figure 15–40**). In muscle cells, Ca^{2+} triggers contraction, and in many secretory cells, including nerve cells, it triggers secretion. Ca^{2+} can act as a signal in this way because its concentration in the cytosol is normally very low (~10^{-7} M), whereas its concentration in the extracellular fluid (~10^{-3} M) and in the lumen of the ER [and sarcoplasmic reticulum (SR) in muscle] is high. Thus, there is a large gradient tending to drive Ca^{2+} into the cytosol across both the plasma membrane and the ER or SR membrane. When a signal transiently opens Ca^{2+} channels in these membranes, Ca^{2+} rushes into the cytosol, increasing the local Ca^{2+} concentration 10–20-fold and activating Ca^{2+}-responsive proteins in the cell. <CGTC>

The Ca^{2+} from outside the cell enters the cytosol through various Ca^{2+} channels in the plasma membrane, which open in response to ligand binding, stretch, or membrane depolarization. The Ca^{2+} from the ER enters the cytosol through either IP_3 receptors (see Figure 15–39) or **ryanodine receptors** (so called because they are sensitive to the plant alkaloid ryanodine). Ryanodine receptors are normally activated by Ca^{2+} binding and thereby amplify the Ca^{2+} signal. Ca^{2+} also activates IP_3 receptors but only in the presence of IP_3; and very high concentrations of Ca^{2+} inactivate them. We discuss how Ca^{2+} is released from the sarcoplasmic reticulum to cause the contraction of muscle cells in Chapter 16.

Several mechanisms keep the concentration of Ca^{2+} in the cytosol low in resting cells (**Figure 15–41**). All eucaryotic cells have a Ca^{2+}-pump in their plasma membrane that uses the energy of ATP hydrolysis to pump Ca^{2+} out of the cytosol. Cells such as muscle and nerve cells, which make extensive use of Ca^{2+} signaling, have an additional Ca^{2+} transport protein (a Na^+-driven Ca^{2+} exchanger) in their plasma membrane that couples the efflux of Ca^{2+} to the influx of Na^+. A Ca^{2+} pump in the ER membrane also has an important role in keeping the cytosolic Ca^{2+} concentration low: this Ca^{2+}-pump enables the ER to take up large amounts of Ca^{2+} from the cytosol against a steep concentration gradient, even when Ca^{2+} levels in the cytosol are low. In addition, a low-affinity, high-capacity Ca^{2+} pump in the inner mitochondrial membrane has an important role in limiting the Ca^{2+} signal and in terminating it; it uses the electrochemical gradient generated across this membrane during the electron-transfer steps of oxidative phosphorylation to take up Ca^{2+} from the cytosol. The resulting increase in Ca^{2+} concentration in the mitochondrion can activate some enzymes of the citric acid cycle, thereby increasing the synthesis of ATP and linking cell activation to energy production; an excessive increase in mitochondrial Ca^{2+}, however, leads to cell death.

The Frequency of Ca^{2+} Oscillations Influences a Cell's Response

Researchers often use Ca^{2+}-sensitive fluorescent indicators, such as *aequorin* or *fura-2* (discussed in Chapter 9), to monitor cytosolic Ca^{2+} in individual cells after activation of the inositol phospholipid signaling pathway. When viewed in this

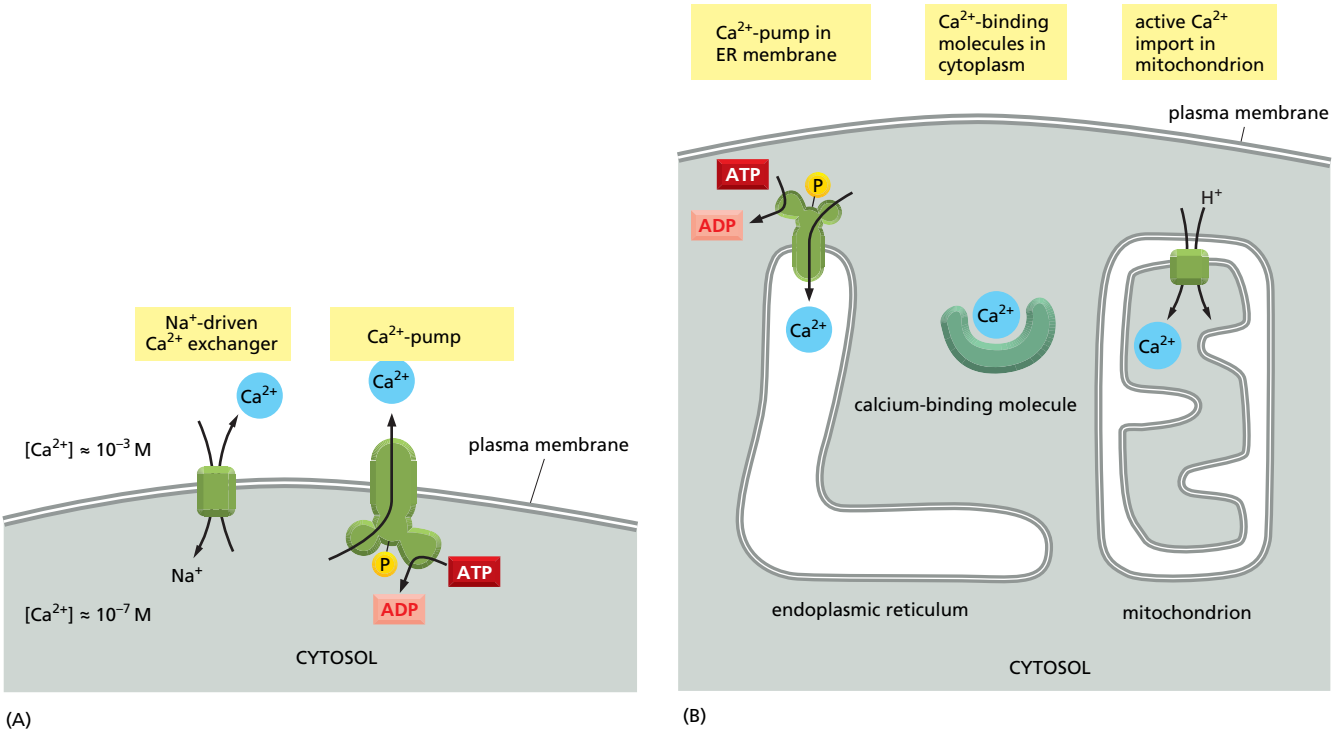


Figure 15–41 The main ways in which eucaryotic cells maintain a very low concentration of free Ca^{2+} in their cytosol. (A) Ca^{2+} is actively pumped out of the cytosol to the cell exterior. (B) Ca^{2+} is pumped out of the cytosol into the ER and mitochondria, and various molecules in the cytosol bind free Ca^{2+} tightly.

way, the initial Ca^{2+} signal appears small and localized to one or more discrete regions of the cell. These Ca^{2+} puffs or sparks reflect the local opening of individual or small groups of Ca^{2+}-release channels in the ER. Because various Ca^{2+}-binding proteins act as Ca^{2+} buffers and restrict the diffusion of Ca^{2+}, the signal often remains localized to the site where the Ca^{2+} entered the cytosol. If the extracellular signal is sufficiently strong and persistent, however, this localized Ca^{2+} signal can propagate as a regenerative Ca^{2+} wave through the cytosol (see Figure 15–40), much like an action potential in an axon. Such a Ca^{2+} "spike" is often followed by a series of further spikes, each usually lasting seconds (**Figure 15–42**). Such *Ca^{2+} oscillations* can persist for as long as receptors are activated at the cell surface. Both the waves and the oscillations are thought to depend, at least in part, on a combination of positive and negative feedback by Ca^{2+} on

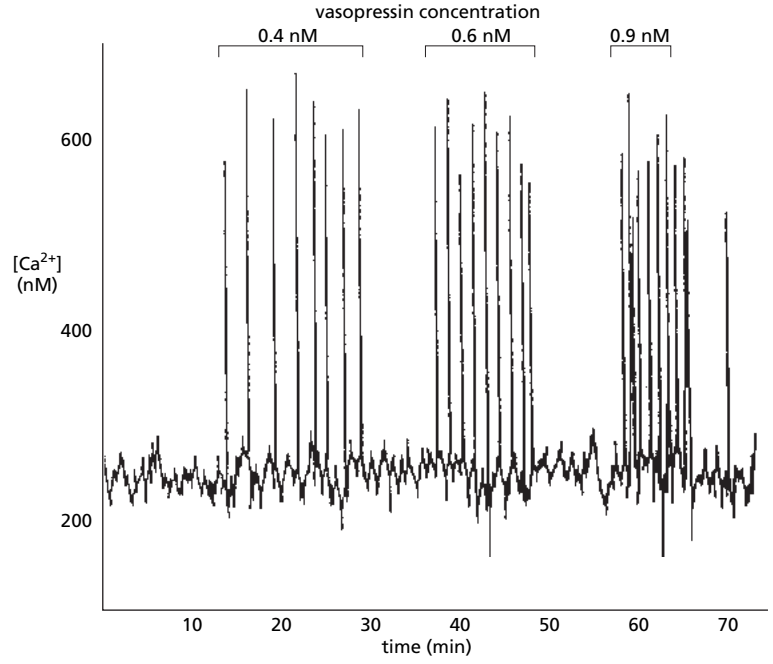


Figure 15–42 Vasopressin-induced Ca^{2+} oscillations in a liver cell. The cell was loaded with the Ca^{2+}-sensitive protein aequorin and then exposed to increasing concentrations of the peptide signal molecule *vasopressin*, which activates a GPCR and thereby PLCβ (see Tabe 15–2). Note that the frequency of the Ca^{2+} spikes increases with an increasing concentration of vasopressin but that the amplitude of the spikes is not affected. Each spike lasts about 7 seconds. (Adapted from N.M. Woods, K.S.R. Cuthbertson and P.H. Cobbold, *Nature* 319:600–602, 1986. With permission from Macmillan Publishers Ltd.)

both the IP_3 receptors and ryanodine receptors: the released Ca^{2+} initially stimulates more Ca^{2+} release from both receptors, a process known as *Ca^{2+}-induced Ca^{2+} release;* but then, as its concentration gets high enough, the Ca^{2+} inhibits further release; and this delayed negative feedback gives rise to oscillations (see Figure 15–28D).

One can follow the effect of the Ca^{2+} oscillations on specific Ca^{2+}-sensitive proteins by using real-time imaging of individual cells expressing fluorescent reporter proteins. One can show, for example, that each Ca^{2+} spike in some signal-induced responses will recruit PKC transiently to the plasma membrane, where it will transiently phosphorylate a reporter protein.

The frequency of the Ca^{2+} oscillations reflects the strength of the extracellular stimulus (see Figure 15–42), and this frequency can be translated into a frequency-dependent cell response. In some cases, the frequency-dependent response itself is also oscillatory: in hormone-secreting pituitary cells, for example, stimulation by an extracellular signal induces repeated Ca^{2+} spikes, each of which is associated with a burst of hormone secretion. In other cases, the frequency-dependent response is non-oscillatory: in some types of cells, for instance, one frequency of Ca^{2+} spikes activates the transcription of one set of genes, while a higher frequency activates the transcription of a different set. How do cells sense the frequency of Ca^{2+} spikes and change their response accordingly? The mechanism presumably depends on Ca^{2+}-sensitive proteins that change their activity as a function of Ca^{2+} spike frequency. A protein kinase that acts as a molecular memory device seems to have this remarkable property, as we discuss next.

Ca^{2+}/Calmodulin-Dependent Protein Kinases (CaM-Kinases) Mediate Many of the Responses to Ca^{2+} Signals in Animal Cells

Various Ca^{2+}-binding proteins help to relay the cytosolic Ca^{2+} signal. The most important is **calmodulin**, which is found in all eucaryotic cells and can constitute as much as 1% of the total protein mass. Calmodulin functions as a multipurpose intracellular Ca^{2+} receptor, governing many Ca^{2+}-regulated processes. It consists of a highly conserved, single polypeptide chain with four high-affinity Ca^{2+}-binding sites (**Figure 15–43**A). When activated by Ca^{2+} binding, it undergoes a conformational change. Because two or more Ca^{2+} ions must bind before calmodulin adopts its active conformation, the protein responds in an almost switchlike manner to increasing concentrations of Ca^{2+} (see Figure 15–25): a tenfold increase in Ca^{2+} concentration typically causes a fiftyfold increase in calmodulin activation.

The allosteric activation of calmodulin by Ca^{2+} is analogous to the allosteric activation of PKA by cyclic AMP, except that Ca^{2+}/calmodulin has no enzymatic

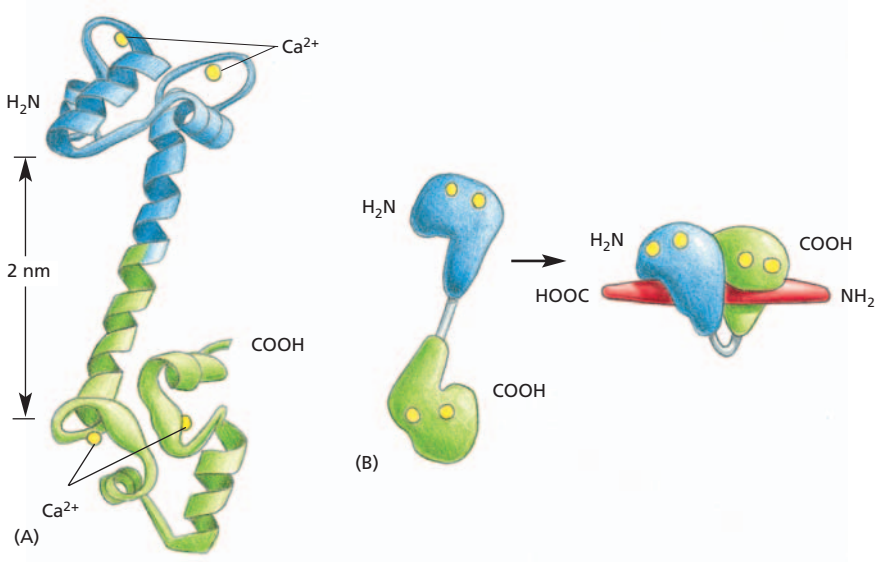


Figure 15–43 The structure of Ca^{2+}/calmodulin based on x-ray diffraction and NMR studies. <CTTC> (A) The molecule has a dumbbell shape, with two globular ends, which can bind to many target proteins. The globular ends are connected by a long, exposed α helix, which allows the protein to adopt a number of very different conformations, depending on the target protein it interacts with. Each globular head has two Ca^{2+}-binding domains. (B) The major structural change in Ca^{2+}/calmodulin that occurs when it binds to a target protein (in this example, a peptide that consists of the Ca^{2+}/calmodulin-binding domain of a Ca^{2+}/calmodulin-dependent protein kinase). Note that the Ca^{2+}/calmodulin has "jack-knifed" to surround the peptide. When it binds to other targets, it can adopt different conformations. (A, based on x-ray crystallographic data from Y.S. Babu et al., *Nature* 315:37–40, 1985. With permission from Macmillan Publishers Ltd; B, based on x-ray crystallographic data from W.E. Meador, A.R. Means and F.A. Quiocho, *Science* 257:1251–1255, 1992, and on NMR data from M. Ikura et al., *Science* 256:632–638, 1992. With permission from AAAS.)

activity itself but instead acts by binding to and activating other proteins. In some cases, calmodulin serves as a permanent regulatory subunit of an enzyme complex, but usually the binding of Ca^{2+} instead enables calmodulin to bind to various target proteins in the cell to alter their activity.

When an activated molecule of Ca^{2+}/calmodulin binds to its target protein, the calmodulin further changes its conformation, the nature of which depends on the specific target protein (Figure 15–43B). Among the many targets calmodulin regulates are enzymes and membrane transport proteins. As one example, Ca^{2+}/calmodulin binds to and activates the plasma membrane Ca^{2+}-pump that uses ATP hydrolysis to pump Ca^{2+} out of cells (see Figure 15–41). Thus, whenever the concentration of Ca^{2+} in the cytosol rises, the pump is activated, which helps to return the cytosolic Ca^{2+} level to resting levels.

Many effects of Ca^{2+}, however, are more indirect and are mediated by protein phosphorylations catalyzed by a family of serine/threonine protein kinases called **Ca^{2+}/calmodulin-dependent kinases (CaM-kinases)**. Some CaM-kinases phosphorylate gene regulatory proteins, such as the CREB protein (see Figure 15–36), and in this way activate or inhibit the transcription of specific genes.

One of the best-studied CaM-kinases is **CaM-kinase II**, which is found in most animal cells but is especially enriched in the nervous system. It constitutes up to 2% of the total protein mass in some regions of the brain, and it is highly concentrated in synapses. CaM-kinase II has two remarkable properties that are related. First, it can function as a molecular memory device, switching to an active state when exposed to Ca^{2+}/calmodulin and then remaining active even after the Ca^{2+} signal has decayed. This is because the kinase phosphorylates itself (a process called *autophosphorylation*), as well as other cell proteins, when Ca^{2+}/calmodulin activates it. In its autophosphorylated state, the enzyme remains active even in the absence of Ca^{2+}, thereby prolonging the duration of the kinase activity beyond that of the initial activating Ca^{2+} signal. The enzyme maintains this activity until serine/threonine protein phosphatases overwhelm the autophosphorylation and shut the kinase off (**Figure 15–44**). CaM-kinase II activation can thereby serve as a memory trace of a prior Ca^{2+} pulse, and it seems to have a role in some types of memory and learning in the vertebrate nervous system. Mutant mice that lack a brain-specific form of the enzyme have specific defects in their ability to remember where things are.

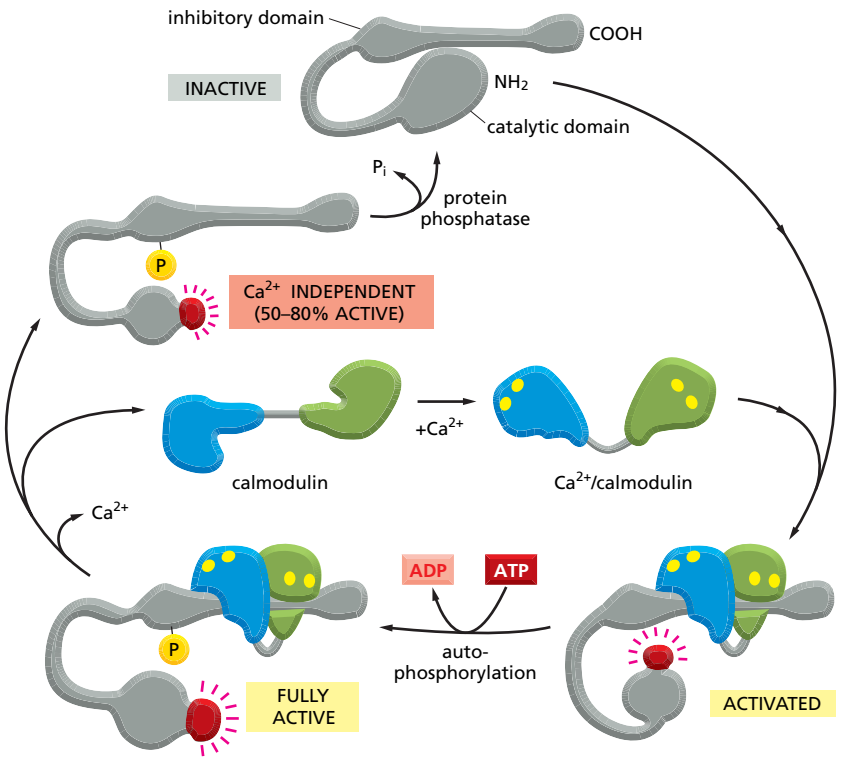


Figure 15–44 The stepwise activation of CaM-kinase II. The enzyme is a large protein complex of 12 subunits, although, for simplicity, only one subunit is shown (in *gray*). In the absence of Ca^{2+}/calmodulin, the enzyme is inactive as the result of an interaction between an inhibitory domain and the catalytic domain. Ca^{2+}/calmodulin binding alters the conformation of the protein, partially activating it. The catalytic domains in the complex phosphorylate the inhibitory domains of neighboring subunits, as well as other proteins in the cell (not shown). The autophosphorylation of the enzyme complex (by mutual phosphorylation of its subunits) fully activates the enzyme. It also prolongs the activity of the enzyme in two ways. First, it traps the bound Ca^{2+}/calmodulin so that it does not dissociate from the enzyme complex until cytosolic Ca^{2+} levels return to basal values for at least 10 seconds (not shown). Second, it converts the enzyme to a Ca^{2+}-independent form so that the kinase remains active even after the Ca^{2+}/calmodulin dissociates from it. This activity continues until the action of a protein phosphatase overrides the autophosphorylation activity of CaM-kinase II.

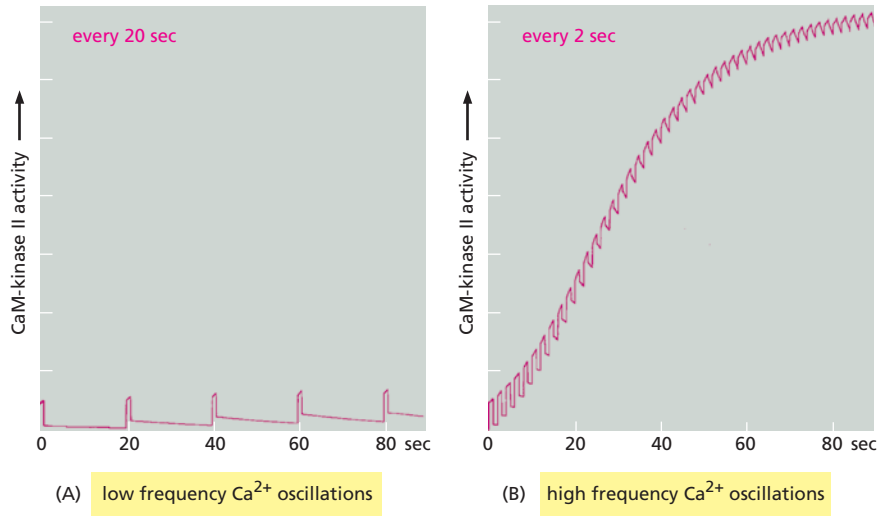


Figure 15–45 CaM-kinase II as a frequency decoder of Ca^{2+} oscillations. (A) At low frequencies of Ca^{2+} spikes, the enzyme becomes inactive after each spike, as the autophosphorylation induced by Ca^{2+}/calmodulin binding does not maintain the enzyme's activity long enough for the enzyme to remain active until the next Ca^{2+} spike arrives. (B) At higher spike frequencies, however, the enzyme fails to inactivate completely between Ca^{2+} spikes, so its activity ratchets up with each spike. If the spike frequency is high enough, this progressive increase in enzyme activity will continue until the enzyme is autophosphorylated on all subunits and is therefore maximally activated. Although not shown, once enough of its subunits are autophosphorylated, the enzyme can be maintained in a highly active state even with a relatively low frequency of Ca^{2+} spikes (a form of cell memory).

The binding of Ca^{2+}/calmodulin to the enzyme is enhanced by the CaM-kinase II autophosphorylation (an additional form of positive feedback), helping the response of the enzyme to repeated Ca^{2+} spikes to exhibit a steep threshold in its frequency response, as discussed earlier. (From P.I. Hanson, T. Meyer, L. Stryer and H. Schulman, *Neuron* 12:943–956, 1994. With permission from Elsevier.)

The second remarkable property of CaM-kinase II follows from the first, in that the enzyme can use its intrinsic memory mechanism to act as a frequency decoder of Ca^{2+} oscillations. This property is thought to be especially important at a nerve cell synapse, where changes in intracellular Ca^{2+} levels in a postsynaptic cell as a result of neural activity can lead to long-term changes in the subsequent effectiveness of that synapse (discussed in Chapter 11). When CaM-kinase II is immobilized on a solid surface and exposed to both a protein phosphatase and repetitive pulses of Ca^{2+}/calmodulin at different frequencies that mimic those observed in stimulated cells, the enzyme's activity increases steeply as a function of pulse frequency (**Figure 15–45**). Moreover, the frequency response of this multisubunit enzyme depends on its exact subunit composition, so that a cell can tailor its response to Ca^{2+} oscillations to particular needs by adjusting the composition of the CaM-kinase II enzyme that it makes.

Some G Proteins Directly Regulate Ion Channels

G proteins do not act exclusively by regulating the activity of membrane-bound enzymes that alter the concentration of cyclic AMP or Ca^{2+} in the cytosol. The α subunit of one type of G protein (called G_{12}), for example, activates a guanine nucleotide exchange factor (GEF) that activates a monomeric GTPase of the *Rho family* (discussed later and in Chapter 16), which regulates the actin cytoskeleton.

In some other cases, G proteins directly activate or inactivate ion channels in the plasma membrane of the target cell, thereby altering the ion permeability—and hence the electrical excitability—of the membrane. As an example, acetylcholine released by the vagus nerve reduces both the rate and strength of heart muscle cell contraction (see Figure 15–9B). This effect is mediated by a special class of acetylcholine receptors that activate the G_i protein discussed earlier. Once activated, the α subunit of G_i inhibits adenylyl cyclase (as described previously), while the βγ subunits bind to K^+ channels in the heart muscle cell plasma membrane and open them. The opening of these K^+ channels makes it harder to depolarize the cell and thereby contributes to the inhibitory effect of acetylcholine on the heart. (These acetylcholine receptors, which can be activated by the fungal alkaloid muscarine, are called *muscarinic acetylcholine receptors* to distinguish them from the very different *nicotinic acetylcholine receptors*, which are ion-channel-coupled receptors on skeletal muscle and nerve cells that can be activated by the binding of nicotine, as well as by acetylcholine.)

Other G proteins regulate the activity of ion channels less directly, either by stimulating channel phosphorylation (by PKA, PKC, or CaM-kinase, for example) or by causing the production or destruction of cyclic nucleotides that directly activate or inactivate ion channels. These *cyclic-nucleotide-gated ion channels* have a crucial role in both smell (olfaction) and vision, as we now discuss.

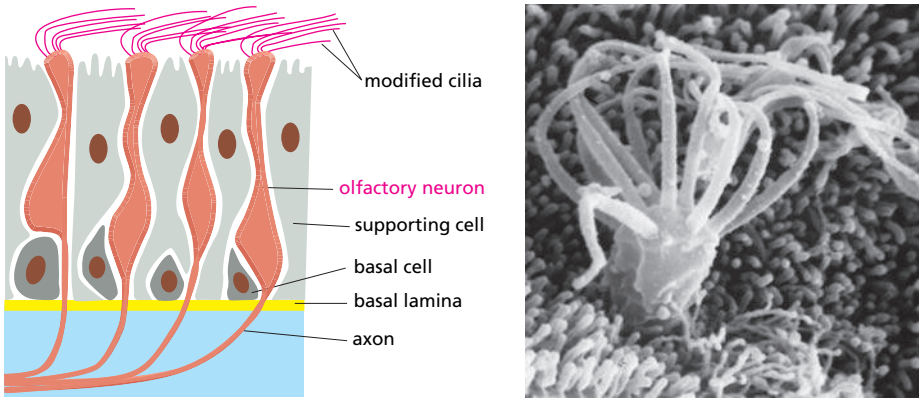


Figure 15–46 Olfactory receptor neurons. (A) A section of olfactory epithelium in the nose. Olfactory receptor neurons possess modified cilia, which project from the surface of the epithelium and contain the olfactory receptors, as well as the signal transduction machinery. The axon, which extends from the opposite end of the receptor neuron, conveys electrical signals to the brain when an odorant activates the cell to produce an action potential. In rodents, at least, the basal cells act as stem cells, producing new receptor neurons throughout life, to replace the neurons that die. (B) A scanning electron micrograph of the cilia on the surface of an olfactory neuron. (B, from E.E. Morrison and R.M. Costanzo, *J. Comp. Neurol.* 297:1–13, 1990. With permission from Wiley-Liss.)

Smell and Vision Depend on GPCRs That Regulate Cyclic-Nucleotide-Gated Ion Channels

Humans can distinguish more than 10,000 distinct smells, which they detect using specialized olfactory receptor neurons in the lining of the nose. These cells use specific GPCRs called **olfactory receptors** to recognize odors; the receptors are displayed on the surface of the modified cilia that extend from each cell (**Figure 15–46**). The receptors act through cyclic AMP. When stimulated by odorant binding, they activate an olfactory-specific G protein (known as G_{olf}), which in turn activates adenylyl cyclase. The resulting increase in cyclic AMP opens *cyclic-AMP-gated cation channels*, thereby allowing an influx of Na^+, which depolarizes the olfactory receptor neuron and initiates a nerve impulse that travels along its axon to the brain.

There are about 1000 different olfactory receptors in a mouse and about 350 in a human, each encoded by a different gene and each recognizing a different set of odorants. Each olfactory receptor neuron produces only one of these receptors (see p. 453); the neuron responds to a specific set of odorants by means of the specific receptor it displays, and each odorant activates its own characteristic set of olfactory receptor neurons. The same receptor also helps direct the elongating axon of each developing olfactory neuron to the specific target neurons that it will connect to in the brain. A different set of GPCRs acts in a similar way to mediate responses to *pheromones,* chemical signals detected in a different part of the nose that are used in communication between members of the same species. Humans, however, lack functional pheromone receptors.

Vertebrate vision employs a similarly elaborate, highly sensitive, signal detection process. Cyclic-nucleotide-gated ion channels are also involved, but the crucial cyclic nucleotide is **cyclic GMP** (**Figure 15–47**) rather than cyclic AMP. As with cyclic AMP, a continuous rapid synthesis (by *guanylyl cyclase*) and rapid degradation (by *cyclic GMP phosphodiesterase*) controls the concentration of cyclic GMP in the cytosol.

In visual transduction responses, which are the fastest G-protein-mediated responses known in vertebrates, the receptor activation stimulated by light causes a fall rather than a rise in the level of the cyclic nucleotide. The pathway has been especially well studied in **rod photoreceptors** (**rods**) in the vertebrate retina. Rods are responsible for noncolor vision in dim light, whereas *cone photoreceptors (cones)* are responsible for color vision in bright light. A rod photoreceptor is a highly specialized cell with outer and inner segments, a cell body, and a synaptic region where the rod passes a chemical signal to a retinal nerve cell (**Figure 15–48**). This nerve cell relays the signal to another nerve cell in the retina, which in turn relays it to the brain (see Figure 23–16).

The phototransduction apparatus is in the outer segment of the rod, which contains a stack of *discs*, each formed by a closed sac of membrane that has many embedded photosensitive **rhodopsin** molecules. The plasma membrane surrounding the outer segment contains *cyclic-GMP-gated cation channels*. Cyclic GMP bound to these channels keeps them open in the dark. Paradoxically,

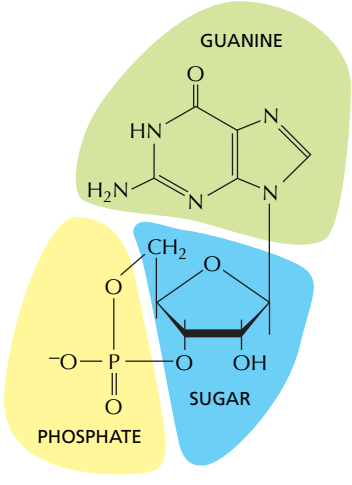
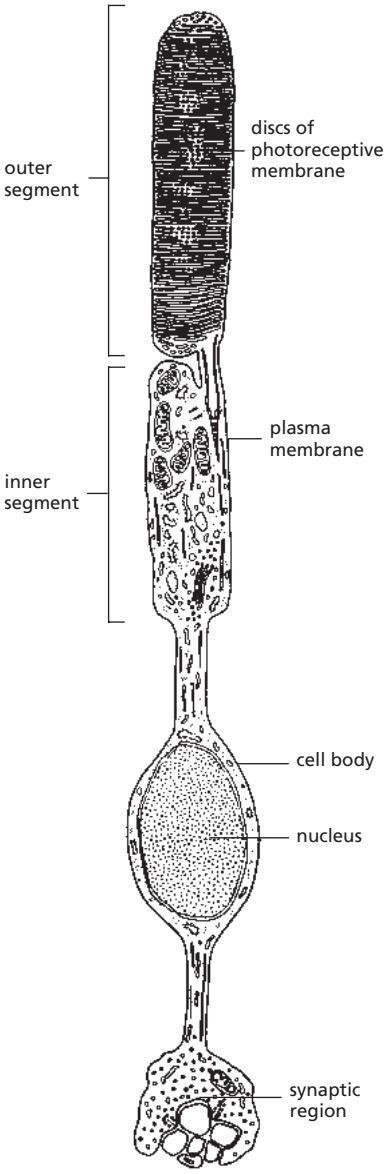


Figure 15–47 Cyclic GMP.

Figure 15–48 A rod photoreceptor cell. There are about 1000 discs in the outer segment. The disc membranes are not connected to the plasma membrane. The inner and outer segments are specialized parts of a *primary cilium* (discussed in Chapter 16); as mentioned earlier and discussed later, a primary cilium extends from the surface of most vertebrate cells, where it serves as a signaling organelle.



light causes a hyperpolarization (which inhibits synaptic signaling) rather than a depolarization of the plasma membrane (which would stimulate synaptic signaling). Hyperpolarization (that is, the membrane potential moves to a more negative value—discussed in Chapter 11) results because the light-induced activation of rhodopsin molecules in the disc membrane decreases the cyclic GMP concentration and closes the cation channels in the surrounding plasma membrane (**Figure 15–49**).

Rhodopsin is a member of the GPCR family, but the activating extracellular signal is not a molecule but a photon of light. Each rhodopsin molecule contains a covalently attached chromophore, 11-*cis* retinal, which isomerizes almost instantaneously to all-*trans* retinal when it absorbs a single photon. The isomerization alters the shape of the retinal, forcing a conformational change in the protein (opsin). The activated rhodopsin molecule then alters the conformation of the G-protein *transducin* (G_t), causing the transducin α subunit to activate **cyclic GMP phosphodiesterase**. The phosphodiesterase then hydrolyzes cyclic GMP, so that cyclic GMP levels in the cytosol fall. This drop in cyclic GMP concentration decreases the amount of cyclic GMP bound to the plasma membrane cation channels, allowing more of these cyclic-GMP-sensitive channels to close. In this way, the signal quickly passes from the disc membrane to the plasma membrane, and a light signal is converted into an electrical one, through a hyperpolarization of the rod cell plasma membrane.

Rods use several negative feedback loops to allow the cells to revert quickly to a resting, dark state in the aftermath of a flash of light—a requirement for perceiving the shortness of the flash. A rhodopsin-specific kinase called *rhodopsin kinase (RK)* phosphorylates the cytosolic tail of activated rhodopsin on multiple serines, partially inhibiting the ability of the rhodopsin to activate transducin. An

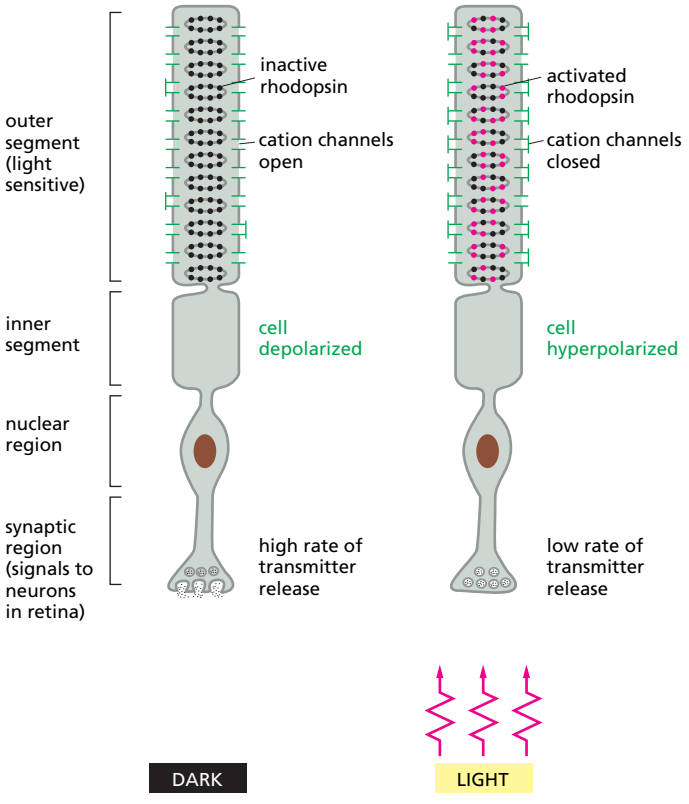


Figure 15–49 The response of a rod photoreceptor cell to light. Rhodopsin molecules in the outer-segment discs absorb photons. Photon absorption closes cation channels in the plasma membrane, which hyperpolarizes the membrane and reduces the rate of neurotransmitter release from the synaptic region. Because the neurotransmitter inhibits many of the postsynaptic retinal neurons, illumination serves to free the neurons from inhibition and thus, in effect, excites them.

Table 15–3 Four Major Families of Trimeric G Proteins*

FAMILY	SOME FAMILY MEMBERS	SUBUNITS THAT MEDIATE ACTION	SOME FUNCTIONS
I	G_s	α	activates adenylyl cyclase; activates Ca^{2+} channels
	G_{olf}	α	activates adenylyl cyclase in olfactory sensory neurons
II	G_i	α	inhibits adenylyl cyclase
		βγ	activates K^+ channels
	G_o	βγ	activates K^+ channels; inactivates Ca^{2+} channels
		α and βγ	activates phospholipase C-β
	G_t (transducin)	α	activates cyclic GMP phosphodiesterase in vertebrate rod photoreceptors
III	G_q	α	activates phospholipase C-β
IV	$G_{12/13}$	α	activates Rho family monomeric GTPases (via Rho-GEF) to regulate the actin cytoskeleton

*Families are determined by amino acid sequence relatedness of the α subunits. Only selected examples are included. About 20 α subunits and at least 6 β subunits and 11 γ subunits have been described in humans.

inhibitory protein called *arrestin* then binds to the phosphorylated rhodopsin, further inhibiting rhodopsin's activity. Mice or humans with a mutation that inactivates the gene encoding RK have a prolonged light response, and their rods eventually die.

At the same time as arrestin shuts off rhodopsin, an RGS protein (see p. 896) binds to activated transducin, stimulating the transducin to hydrolyze its bound GTP to GDP, which returns transducin to its inactive state. In addition, the cation channels that close in response to light are permeable to Ca^{2+}, as well as to Na^+, so that when they close, the normal influx of Ca^{2+} is inhibited, causing the Ca^{2+} concentration in the cytosol to fall. The decrease in Ca^{2+} concentration stimulates guanylyl cyclase to replenish the cyclic GMP, rapidly returning its level to where it was before the light was switched on. A specific Ca^{2+}-sensitive protein mediates the activation of guanylyl cyclase in response to a fall in Ca^{2+} levels. In contrast to calmodulin, this protein is inactive when Ca^{2+} is bound to it and active when it is Ca^{2+}-free. It therefore stimulates the cyclase when Ca^{2+} levels fall following a light response.

Negative feedback mechanisms do more than just return the rod to its resting state after a transient light flash; they also help the rod to *adapt,* stepping down the response when the rod is exposed to light continuously. Adaptation, as we discussed earlier, allows the receptor cell to function as a sensitive detector of *changes* in stimulus intensity over an enormously wide range of baseline levels of stimulation. It is why we can see a camera flash in broad daylight.

The various trimeric G proteins we have discussed in this chapter fall into four major families, as summarized in **Table 15–3**.

Intracellular Mediators and Enzymatic Cascades Amplify Extracellular Signals

Despite the differences in molecular details, the different intracellular signaling pathways that GPCRs trigger share certain features and obey similar general principles. They depend on relay chains of intracellular signaling proteins and small intracellular mediators. In contrast to the more direct signaling pathways used by nuclear receptors discussed earlier, these relay chains provide numerous opportunities for amplifying the responses to extracellular signals. In the visual transduction cascade, for example, a single activated rhodopsin molecule catalyzes the activation of hundreds of molecules of transducin at a rate of about 1000 transducin molecules per second. Each activated transducin molecule activates a molecule of cyclic GMP phosphodiesterase, each of which hydrolyzes about 4000 molecules of cyclic GMP per second. This catalytic cascade lasts for about 1 second and results in the hydrolysis of more than 10^5 cyclic GMP molecules for a single quantum of light absorbed, and the resulting

drop in the concentration of cyclic GMP in turn transiently closes hundreds of cation channels in the plasma membrane (**Figure 15–50**). As a result, a rod cell can respond to even a single photon of light in a way that is highly reproducible in its timing and magnitude.

Likewise, when an extracellular signal molecule binds to a receptor that indirectly activates adenylyl cyclase via G_s, each receptor protein may activate many molecules of G_s protein, each of which can activate a cyclase molecule. Each cyclase molecule, in turn, can catalyze the conversion of a large number of ATP molecules to cyclic AMP molecules. A similar amplification operates in the inositol phospholipid signaling pathway. In these ways, a nanomolar (10^{-9} M) change in the concentration of an extracellular signal can induce micromolar (10^{-6} M) changes in the concentration of a small intracellular mediator such as cyclic AMP or Ca^{2+}. Because these mediators function as allosteric effectors to activate specific enzymes or ion channels, a single extracellular signal molecule can alter many thousands of protein molecules within the target cell.

Any such amplifying cascade of stimulatory signals requires counterbalancing mechanisms at every step of the cascade to restore the system to its resting state when stimulation ceases. As emphasized earlier, the response to stimulation can be rapid only if the inactivating mechanisms are also rapid. Cells therefore have efficient mechanisms for rapidly degrading (and resynthesizing) cyclic nucleotides and for buffering and removing cytosolic Ca^{2+}, as well as for inactivating the responding enzymes and ion channels once they have been activated. This is not only essential for turning a response off, it is also important for defining the resting state from which a response begins.

Each protein in the signaling relay chain can be a separate target for regulation, including the receptor itself, as we discuss next.

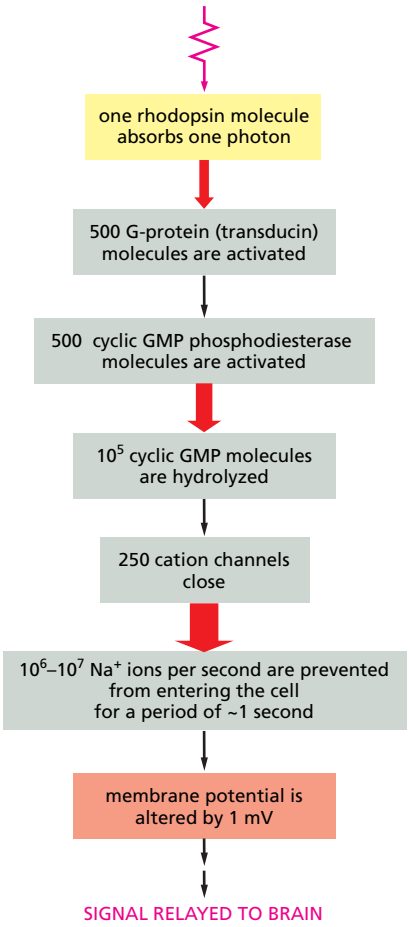


Figure 15–50 Amplification in the light-induced catalytic cascade in vertebrate rods. The *red arrows* indicate the steps where amplification occurs, with the thickness of the arrow roughly indicating the magnitude of the amplification.

GPCR Desensitization Depends on Receptor Phosphorylation

When target cells are exposed to a high concentration of a stimulating ligand for a prolonged period, they can become *desensitized*, or *adapted*, in several different ways. An important class of desensitization mechanisms depend on alteration of the quantity or condition of the receptor molecules themselves.

For GPCRs, there are three general modes of desensitization (see Figure 15–29): (1) In *receptor inactivation*, they become altered so that they can no longer interact with G proteins. (2) In *receptor sequestration*, they are temporarily moved to the interior of the cell (internalized) so that they no longer have access to their ligand. (3) In *receptor down-regulation*, they are destroyed in lysosomes after internalization.

In each case, the desensitization of the GPCRs depends on their phosphorylation by PKA, PKC, or a member of the family of **GPCR kinases** (**GRKs**), which includes the rhodopsin-specific kinase RK involved in rod photoreceptor desensitization discussed earlier. The GRKs phosphorylate multiple serines and threonines on a GPCR, but they do so only after ligand binding has activated the receptor, because it is the activated receptor that allosterically activates the GRK. As with rhodopsin, once a receptor has been phosphorylated by a GRK, it binds with high affinity to a member of the **arrestin** family of proteins (**Figure 15–51**).

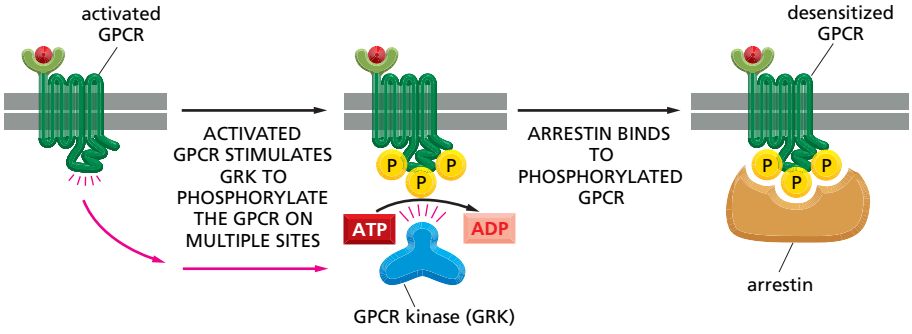


Figure 15–51 The roles of GPCR kinases (GRKs) and arrestins in GPCR desensitization. A GRK phosphorylates only activated receptors because it is the activated GPCR that activates the GRK. The binding of an arrestin to the phosphorylated receptor prevents the receptor from binding to its G protein and also directs its endocytosis (not shown). Mice that are deficient in one form of arrestin fail to desensitize in response to morphine, for example, attesting to the importance of arrestins for desensitization.

The bound arrestin can contribute to the desensitization process in at least two ways. First, it prevents the activated receptor from interacting with G proteins. Second, it serves as an adaptor protein to help couple the receptor to the clathrin-dependent endocytosis machinery (discussed in Chapter 13), inducing receptor-mediated endocytosis. The fate of the internalized GPCR–arrestin complex depends on other proteins in the complex. In some cases, the receptor is dephosphorylated and recycled back to the plasma membrane for reuse. In others, it is ubiquitylated and degraded in lysosomes (discussed later).

Receptor endocytosis does not necessarily stop the receptor from signaling. In some cases, the bound arrestin recruits other signaling proteins to relay the signal onward along new pathways from the internalized GPCRs.

Summary

GPCRs can indirectly activate or inactivate either plasma-membrane-bound enzymes or ion channels via G proteins. When an activated receptor stimulates a G protein, the G protein undergoes a conformational change that activates both its α subunit and its βγ subunits, which can then directly regulate the activity of target proteins in the plasma membrane. Some GPCRs either activate or inactivate adenylyl cyclase, thereby altering the intracellular concentration of the small intracellular mediator cyclic AMP. Others activate a phosphoinositide-specific phospholipase C (PLCβ), which hydrolyzes phosphatidylinositol 4,5-bisphosphate [PI(4,5)P_2] to generate two small intracellular mediators. One is inositol 1,4,5-trisphosphate (IP_3), which releases Ca^{2+} from the ER and thereby increases the concentration of Ca^{2+} in the cytosol. The other is diacylglycerol, which remains in the plasma membrane and helps activate protein kinase C (PKC). An increase in cytosolic cyclic AMP or Ca^{2+} levels affects cells mainly by stimulating protein kinase A (PKA) and Ca^{2+}/calmodulin-dependent protein kinases (CaM-kinases), respectively.

PKC, PKA, and CaM-kinases phosphorylate specific target proteins on serines or threonines and thereby alter the activity of the proteins. Each type of cell has its own characteristic set of target proteins that is regulated in these ways, enabling the cell to make its own distinctive responses to the small intracellular mediators. The intracellular signaling cascades activated by GPCRs greatly amplify the responses, so that many thousands of target protein molecules are changed for each molecule of extracellular signaling ligand bound to its receptor.

The responses mediated by GPCRs are rapidly turned off when the extracellular signal is removed. Thus, the G-protein α subunit is stimulated by its target protein or an RGS to inactivate itself by hydrolyzing its bound GTP to GDP, IP_3 is rapidly dephosphorylated by a lipid phosphatase or phosphorylated by a lipid kinase, cyclic nucleotides are hydrolyzed by phosphodiesterases, Ca^{2+} is rapidly pumped out of the cytosol, and phosphorylated proteins are dephosphorylated by protein phosphatases. Activated GPCRs themselves are phosphorylated by GRKs, thereby trigging arrestin binding. The bound arrestin uncouples the receptors from G proteins and promotes their endocytosis, resulting in either desensitization or continued signaling via arrestin-recruited signaling proteins.

SIGNALING THROUGH ENZYME-COUPLED CELL-SURFACE RECEPTORS

Like GPCRs, **enzyme-coupled receptors** are transmembrane proteins with their ligand-binding domain on the outer surface of the plasma membrane. Instead of having a cytosolic domain that associates with a trimeric G protein, however, their cytosolic domain either has intrinsic enzyme activity or associates directly with an enzyme. Whereas a GPCR has seven transmembrane segments, each subunit of an enzyme-coupled receptor usually has only one. GPCRs and enzyme-coupled receptors often activate some of the same signaling pathways, and there is usually no obvious reason why a particular extracellular signal utilizes one class of receptors rather than the other.

There are six principal classes of enzyme-coupled receptors:

1. *Receptor tyrosine kinases* directly phosphorylate specific tyrosines on themselves and on a small set of intracellular signaling proteins.
2. *Tyrosine-kinase-associated receptors* have no intrinsic enzyme activity but directly recruit cytoplasmic tyrosine kinases to relay the signal.
3. *Receptor serine/threonine kinases* directly phosphorylate specific serines or threonines on themselves and on latent gene regulatory proteins with which they are associated.
4. *Histidine-kinase-associated receptors* activate a two-component signaling pathway in which the kinase phosphorylates itself on histidine and then immediately transfers the phosphoryl group to a second intracellular signaling protein.
5. *Receptor guanylyl cyclases* directly catalyze the production of cyclic GMP in the cytosol, which acts as a small intracellular mediator in much the same way as cyclic AMP.
6. *Receptorlike tyrosine phosphatases* remove phosphate groups from tyrosines of specific intracellular signaling proteins. (They are called "receptorlike" because their presumptive ligands have not yet been identified, and so their receptor function is unproven.)

We focus our discussion on the first four classes, beginning with the receptor tyrosine kinases, the most numerous of the enzyme-coupled receptors.

Activated Receptor Tyrosine Kinases (RTKs) Phosphorylate Themselves

Many extracellular signal proteins act through **receptor tyrosine kinases** (**RTKs**). Notable examples discussed elsewhere in this book include *epidermal growth factor (EGF), platelet-derived growth factor (PDGF), fibroblast growth factors (FGFs), hepatocyte growth factor (HGF), insulin, insulinlike growth factor-1 (IGF1), vascular endothelial growth factor (VEGF), macrophage-colony-stimulating factor (MCSF),* and the *neurotrophins,* including *nerve growth factor (NGF).*

Many cell-surface-bound extracellular signal proteins also act through RTKs. **Ephrins** are the largest class of such membrane-bound ligands, with eight identified in humans. Among their many functions, they help guide the migration of cells and axons along specific pathways during animal development by stimulating responses that result in attraction or repulsion (discussed later and in Chapter 22). The receptors for ephrins, called **Eph receptors**, are also among the most numerous RTKs, with thirteen genes encoding them in humans. The ephrins and Eph receptors are unusual in that they can simultaneously act as both ligand and receptor: on binding to an Eph receptor, some ephrins not only activate the Eph receptor but also become activated themselves to transmit signals into the interior of the ephrin-expressing cell; in this way, the behavior of both the signaling cell and the target cell is altered. Such *bidirectional signaling* between ephrins and Eph receptors is required, for example, to keep some neighboring groups of cells from mixing with each other during development.

There are about 60 genes encoding human RTKs. These receptors can be classified into more than 16 structural subfamilies, each dedicated to its complementary family of protein ligands. **Figure 15–52** shows a number of the families that operate in mammals, and **Table 15–4** lists some of their ligands and functions. In all cases, the binding of the signal protein to the ligand-binding domain on the outside of the cell enables the intracellular tyrosine kinase domain to phosphorylate selected tyrosine side chains, both on the receptor proteins themselves and on intracellular signaling proteins that subsequently bind to the phosphorylated tyrosines on the receptors.

How does the binding of an extracellular ligand activate the kinase domain on the other side of the plasma membrane? For a GPCR, ligand binding is thought to change the relative orientation of several of the transmembrane α helices, thereby shifting the position of the cytoplasmic loops relative to one another. It is difficult to imagine, however, how a conformational change could propagate across the lipid bilayer through a single transmembrane α helix.

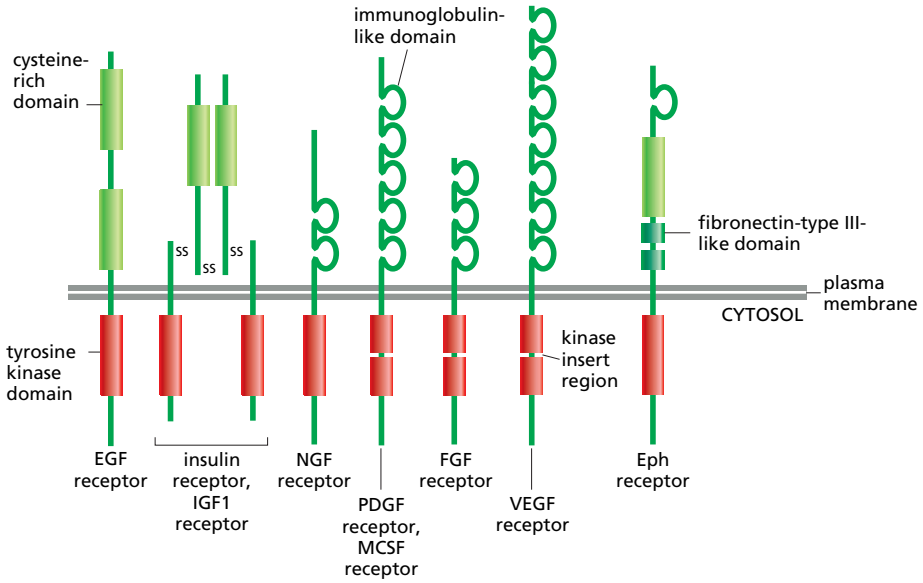


Figure 15–52 Some subfamilies of RTKs. Only one or two members of each subfamily are indicated. Note that the tyrosine kinase domain is interrupted by a "kinase insert region" in some of the subfamilies. The functional roles of most of the cysteine-rich, immunoglobulin-like, and fibronectin-type III-like domains are not known. Some of the ligands for the receptors shown are listed in Table 15–4, along with some representative responses that they mediate.

Instead, for many RTKs, ligand binding causes the receptor chains to dimerize, bringing the kinase domains of two receptor chains together (an example of *induced proximity*, discussed earlier) so that they can become activated and cross-phosphorylate each other on multiple tyrosines, a process referred to as *transautophosphorylation* (**Figure 15–53**A).

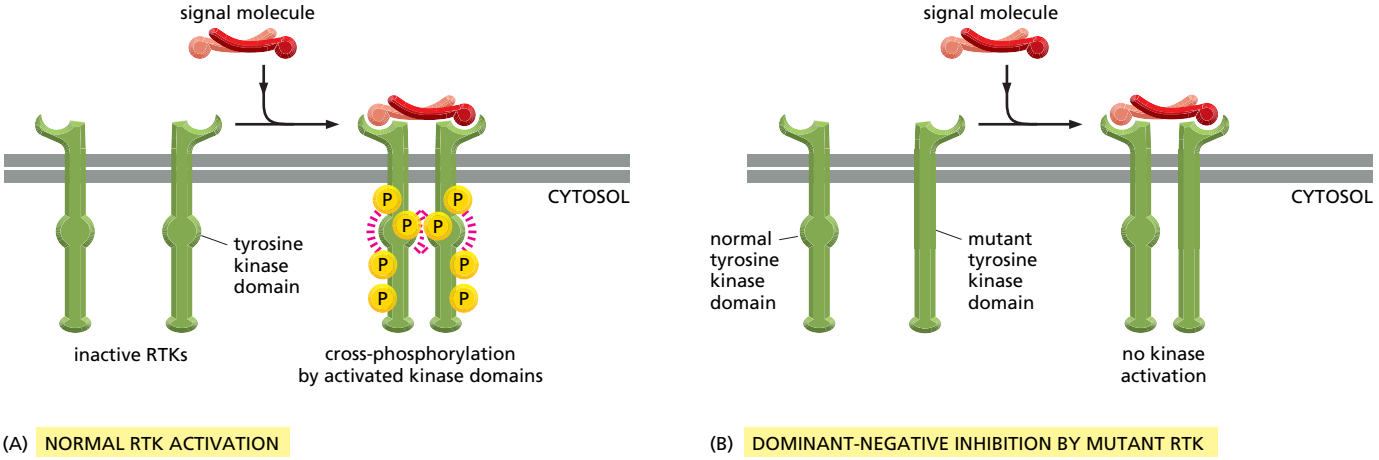
Because of the requirement for receptor dimerization, it is relatively easy to inactivate a specific RTK to determine its importance for a cell response. For this purpose, cells are transfected with DNA encoding a mutant form of the receptor that dimerizes normally but has an inactive kinase domain. When coexpressed at a high level with normal receptors, the mutant receptor acts in a *dominant-negative* way, disabling the normal receptors by forming inactive dimers with them (Figure 15–53B).

Phosphorylated Tyrosines on RTKs Serve as Docking Sites for Intracellular Signaling Proteins

The cross-phosphorylation of adjacent cytosolic tails of RTKs contributes to the receptor activation process in two ways. First, phosphorylation of tyrosines within the kinase domain increases the kinase activity of the enzyme. Second,

Table 15–4 Some Signal Proteins That Act Via RTKs

SIGNAL PROTEIN	RECEPTORS	SOME REPRESENTATIVE RESPONSES
Epidermal growth factor (EGF)	EGF receptors	stimulates cell survival, growth, proliferation, or differentiation of various cell types; acts as inductive signal in development
Insulin	insulin receptor	stimulates carbohydrate utilization and protein synthesis
Insulin-like growth factors (IGF1 and IGF2)	IGF receptor-1	stimulate cell growth and survival in many cell types
Nerve growth factor (NGF)	Trk A	stimulates survival and growth of some neurons
Platelet-derived growth factors (PDGF AA, BB, AB)	PDGF receptors (α and β)	stimulate survival, growth, proliferation, and migration of various cell types
Macrophage-colony-stimulating factor (MCSF)	MCSF receptor	stimulates monocyte/macrophage proliferation and differentiation
Fibroblast growth factors (FGF1 to FGF24)	FGF receptors (FGFR1–FGFR4, plus multiple isoforms of each)	stimulate proliferation of various cell types; inhibit differentiation of some precursor cells; act as inductive signals in development
Vascular endothelial growth factor (VEGF)	VEGF receptors	stimulates angiogenesis
Ephrins (A and B types)	Eph receptors (A and B types)	stimulate angiogenesis; guide cell and axon migration



phosphorylation of tyrosines outside the kinase domain creates high-affinity docking sites for the binding of specific intracellular signaling proteins. Each of the signaling proteins binds to specific phosphorylated sites on the activated receptors because it contains a specific phosphotyrosine-binding domain that recognizes surrounding features of the polypeptide chain in addition to the phosphotyrosine.

Once bound to the activated RTK, a signaling protein may itself become phosphorylated on tyrosines and thereby be activated. In many cases, however, the binding alone may be sufficient to activate the docked signaling protein, by either inducing a conformational change in the protein or simply bringing it near the protein that is next in the signaling pathway. Thus, transautophosphorylation serves as a switch to trigger the transient assembly of an intracellular signaling complex, which can then relay the signal onward, often along multiple routes, to various destinations in the cell (**Figure 15–54**). Because different RTKs bind different combinations of these signaling proteins, they activate different responses.

The receptors for insulin and IGF1 act in a slightly different way. They are tetramers (see Figure 15–52), and ligand binding is thought to rearrange their transmembrane receptor chains, moving the two kinase domains close together. Moreover, most of the phosphotyrosine docking sites generated by ligand binding are not on the receptor itself but on a specialized docking protein called *insulin receptor substrate-1 (IRS1)*. The activated receptor first transautophosphorylates its kinase domains, which then phosphorylate IRS1 on multiple tyrosines, thereby creating many more docking sites than could be accommodated on the receptor alone (see Figure 15–22). Some other RTKs use docking proteins in a similar way to enlarge the size of the signaling complex.

Figure 15–53 Activation and inactivation of RTKs by dimerization. (A) The normal receptors dimerize in response to ligand binding. The two kinase domains cross-phosphorylate each other, increasing the activity of the kinase domains, which now phosphorylate other sites on the receptors. (B) The mutant receptor with an inactivated kinase domain can dimerize normally, but it cannot cross-phosphorylate a normal receptor in a dimer. For this reason, the mutant receptors, if present in excess, will block signaling by the normal receptors—a process called *dominant-negative regulation*. Cell biologists frequently use this strategy to inhibit a specific type of RTK in a cell to determine its normal function. A similar approach can be used to inhibit the function of various other types of receptors and intracellular signaling proteins that function as dimers or larger oligomers.

Proteins with SH2 Domains Bind to Phosphorylated Tyrosines

A whole menagerie of intracellular signaling proteins can bind to the phosphotyrosines on activated RTKs (or on docking proteins such as IRS1). They help to

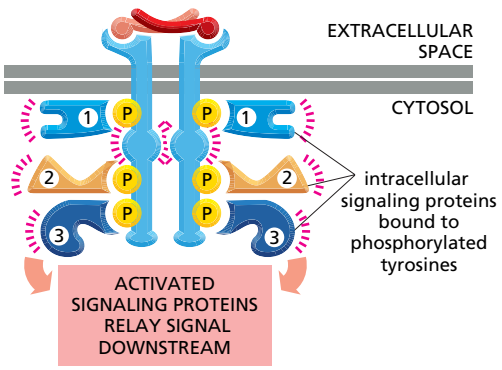


Figure 15–54 The docking of intracellular signaling proteins on phosphotyrosines on an activated RTK. The activated receptor and its bound signaling proteins form a signaling complex that can then broadcast signals along multiple signaling pathways.

relay the signal onward, mainly through chains of protein–protein interactions mediated by modular *interaction domains*, as discussed earlier. Some of the docked proteins are enzymes, such as **phospholipase C-γ (PLCγ)**, which functions in the same way as phospholipase C-β—activating the inositol phospholipid signaling pathway discussed earlier in connection with GPCRs (see Figures 15–38 and 15–39). Through this pathway, RTKs can increase cytosolic Ca^{2+} levels and activate PKC. Another enzyme that docks on these receptors is the cytoplasmic tyrosine kinase *Src*, which phosphorylates other signaling proteins on tyrosines. Yet another is *phosphoinositide 3-kinase (PI 3-kinase)*, which phosphorylates mainly lipids rather than proteins; as we discuss later, the phosphorylated lipids then serve as docking sites to attract various signaling proteins to the plasma membrane.

The intracellular signaling proteins that bind to phosphotyrosines, either on activated RTKs or on proteins docked on them, have varied structures and functions. However, they usually share highly conserved phosphotyrosine-binding domains. These can be either **SH2 domains** (for *Src homology region*) or, less commonly, *PTB domains* (for *phosphotyrosine-binding*). By recognizing specific phosphorylated tyrosines, these small interaction domains enable the proteins that contain them to bind to activated RTKs, as well as to many other intracellular signaling proteins that have been transiently phosphorylated on tyrosines (**Figure 15–55**). As previously discussed, many signaling proteins also contain other interaction domains that allow them to interact specifically with other proteins as part of the signaling process. These domains include the *SH3 domain*, which binds to proline-rich motifs in intracellular proteins (see Figure 15–22). There are about 115 SH2 domains and about 295 SH3 domains encoded in the human genome.

Not all proteins that bind to activated RTKs via SH2 domains help to relay the signal onward. Some act to decrease the signaling process, providing negative

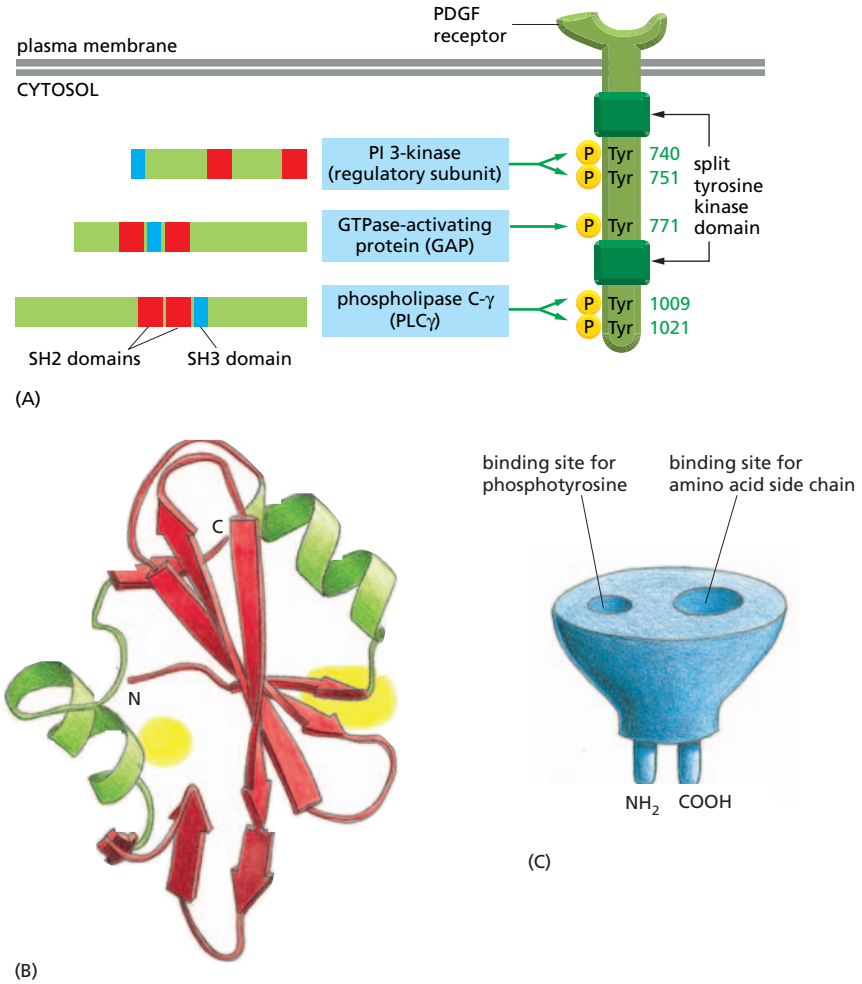


Figure 15–55 The binding of SH2-containing intracellular signaling proteins to an activated PDGF receptor. (A) This drawing of a PDGF receptor shows five phosphotyrosine docking sites, three in the kinase insert region and two on the C-terminal tail, to which the three signaling proteins shown bind as indicated. The numbers on the right indicate the positions of the tyrosines in the polypeptide chain. These binding sites have been identified by using recombinant DNA technology to mutate specific tyrosines in the receptor. Mutation of tyrosines 1009 and 1021, for example, prevents the binding and activation of PLCγ, so that receptor activation no longer stimulates the inositol phospholipid signaling pathway. The locations of the SH2 *(red)* and SH3 *(blue)* domains in the three signaling proteins are indicated. (Additional phosphotyrosine docking sites on this receptor are not shown, including those that bind the cytoplasmic tyrosine kinase Src and two adaptor proteins.) It is unclear how many signaling proteins can bind simultaneously to a single RTK. (B) The three-dimensional structure of an SH2 domain, as determined by x-ray crystallography. The binding pocket for phosphotyrosine is shown in *yellow* on the right, and a pocket for binding a specific amino acid side chain (isoleucine, in this case) is shown in *yellow* on the left (see also Figure 3–39). (C) The SH2 domain is a compact, "plug-in" module, which can be inserted almost anywhere in a protein without disturbing the protein's folding or function (discussed in Chapter 3). Because each domain has distinct sites for recognizing phosphotyrosine and for recognizing a particular amino acid side chain, different SH2 domains recognize phosphotyrosine in the context of different flanking amino acid sequences. (B, based on data from G. Waksman et al., *Cell* 72:779–790, 1993. With permission from Elsevier.)

feedback. One example is the *c-Cbl protein,* which can dock on some activated receptors and catalyze their ubiquitylation, covalently adding a single ubiquitin molecule to one or more sites on the receptor (called *monoubiquityl-ation* to distinguish it from *polyubiquitylation*, in which one or more long ubiquitin chains are added to a protein). Monoubiquitylation promotes the endocytosis and degradation of the receptors in lysosomes—an example of receptor down-regulation (see Figure 15–29). Endocytic proteins that contain *ubiquitin-interaction motifs (UIMs)* recognize the monoubiquitylated RTKs and direct them into clathrin-coated vesicles and, ultimately, into lysosomes (discussed in Chapter 13). Mutations that inactivate c-Cbl-dependent RTK down-regulation cause prolonged RTK signaling and thereby promote the development of cancer.

As is the case for GPCRs, ligand-induced endocytosis of RTKs does not always decrease signaling. In some cases, RTKs are endocytosed with their bound signaling proteins and continue to signal from endosomes or other intracellular compartments. This mechanism, for example, allows *nerve growth factor (NGF)* to bind to its specific RTK (called TrkA) at the end of a long nerve cell axon and signal to the cell body of the same cell a long distance away. Here, signaling endocytic vesicles containing TrkA, with NGF bound on the inside and signaling proteins docked on the cytosolic side, are transported along the axon to the cell body, where they signal the cell to survive.

Some signaling proteins are composed almost entirely of SH2 and SH3 domains and function as *adaptors* to couple tyrosine-phosphorylated proteins to other proteins that do not have their own SH2 domains (see Figure 15–22). Adaptor proteins of this type help to couple activated RTKs to the important signaling protein *Ras*, a monomeric GTPase that, in turn, can activate various downstream signaling pathways, as we now discuss.

Ras Belongs to a Large Superfamily of Monomeric GTPases

The **Ras superfamily** consists of various families of monomeric GTPases, but only the Ras and Rho families relay signals from cell-surface receptors (**Table 15–5**). By interacting with different intracellular signaling proteins, a single Ras or Rho family member can coordinately spread the signal along several distinct downstream signaling pathways, thereby acting as a *signaling hub.*

There are three major, closely related Ras proteins in humans (H-, K-, and N-Ras—see Table 15–5). Although they have subtly different functions, they are thought to work in the same way, and we will refer to them simply as **Ras**. Like many monomeric GTPases, Ras contains one or more covalently attached lipid groups that help anchor the protein to the cytoplasmic face of the membrane where the protein functions, which is mainly the plasma membrane, from where it relays signals to other parts of the cell. Ras is often required, for example, when RTKs signal to the nucleus to stimulate cell proliferation or differentiation, both of which require changes in gene expression. If one inhibits Ras function by the microinjection of either neutralizing anti-Ras antibodies or a dominant-negative mutant form of Ras, the cell proliferation or differentiation

Table 15–5 The Ras Superfamily of Monomeric GTPases

FAMILY	SOME FAMILY MEMBERS	SOME FUNCTIONS
Ras	H-Ras, K-Ras, N-Ras	relay signals from RTKs
	Rheb	activates mTOR to stimulate cell growth
	Rep1	activated by a cyclic-AMP-dependent GEF; influences cell adhesion by activating integrins
Rho*	Rho, Rac, Cdc42	relay signals from surface receptors to the cytoskeleton and elsewhere
ARF*	ARF1–ARF6	regulate assembly of protein coats on intracellular vesicles
Rab*	Rab1–60	regulate intracellular vesicle traffic
Ran*	Ran	regulates mitotic spindle assembly and nuclear transport of RNAs and proteins

*The Rho family is discussed in Chapter 16, the ARF and Rab proteins in Chapter 13, and Ran in Chapters 12 and 17. The three-dimensional structure of Ras is shown in Figure 3–72.

responses normally induced by the activated RTKs do not occur. Conversely, 30% of human tumors have hyperactive mutant forms of Ras, which contribute to the uncontrolled proliferation of the cancer cells.

Like other GTP-binding proteins, Ras functions as a molecular switch, cycling between two distinct conformational states—active when GTP is bound and inactive when GDP is bound (see Figure 15–18B). <GAAC> As discussed earlier for monomeric GTPases in general, two classes of signaling proteins regulate Ras activity by influencing its transition between active and inactive states (see Figure 15–19). *Ras guanine nucleotide exchange factors* (**Ras-GEFs**) stimulate the dissociation of GDP and the subsequent uptake of GTP from the cytosol, thereby activating Ras. *Ras GTPase-activating proteins* (**Ras-GAPs**) increase the rate of hydrolysis of bound GTP by Ras, thereby inactivating Ras. Hyperactive mutant forms of Ras are resistant to GAP-mediated GTPase stimulation and are locked permanently in the GTP-bound active state, which is why they promote the development of cancer.

But how do RTKs normally activate Ras? In principle, they could either activate a Ras-GEF or inhibit a Ras-GAP. Even though some GAPs bind directly (via their SH2 domains) to activated RTKs (see Figure 15–55A), whereas GEFs usually bind only indirectly, it is the indirect coupling of the receptor to a Ras-GEF that drives Ras into its active state. In fact, the loss of function of a Ras-GEF has a similar effect to the loss of function of Ras itself. Activation of the other Ras superfamily proteins, including those of the Rho family, also occurs through the activation of GEFs. The particular GEF determines in which membrane the GTPase is activated and, by acting as a scaffold, which downstream proteins the GTPase activates.

Ras proteins and the proteins that regulate them have been highly conserved in evolution, and genetic analyses in *Drosophila* and *C. elegans* provided the first clues to how RTKs activate Ras. Genetic studies of photoreceptor cell development in the *Drosophila* eye were particularly informative.

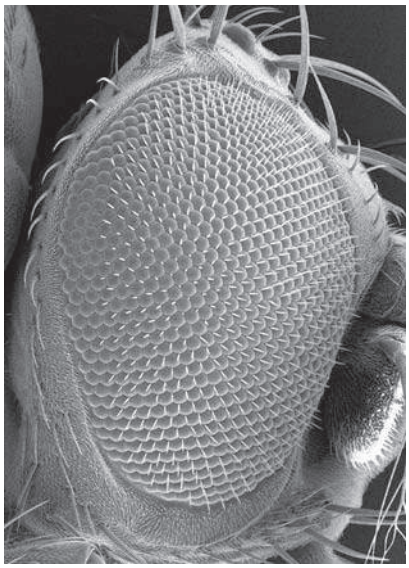


Figure 15–56 Scanning electron micrograph of a compound eye of *Drosophila*. The eye is composed of about 800 identical units (ommatidia), each having a separate lens that focuses light onto eight photoreceptor cells at its base. (Courtesy of Kevin Moses.)

RTKs Activate Ras Via Adaptors and GEFs: Evidence from the Developing *Drosophila* Eye

The *Drosophila* compound eye consists of about 800 identical units called *ommatidia,* each composed of 8 photoreceptor cells (R1–R8) and 12 accessory cells (**Figure 15–56**). The eye develops from a simple epithelial sheet, and the cells that make up each ommatidium are recruited from the sheet in a fixed sequence, by a series of cell–cell interactions. Beginning with the development of the R8 photoreceptor, each differentiating cell induces its uncommitted immediate neighbors to adopt a specific fate and assemble into the developing ommatidium (**Figure 15–57**).

The development of the R7 photoreceptor, which is required for the detection of ultraviolet light, has been studied most intensively, beginning with the description of a mutant fly called *Sevenless (Sev),* in which a deficiency of R7 is the only observed defect. Such mutants are easy to select on the basis of their blindness to ultraviolet light. The normal *Sev* gene was shown to encode an RTK that is expressed on R7 precursor cells. Further genetic analysis of mutants in which R7 development is blocked but the Sev protein itself is not affected led to the identification of the gene *Bride-of-sevenless (Boss),* which encodes the ligand for the Sev RTK. Boss is a seven-pass transmembrane protein that is expressed exclusively on the surface of the adjacent R8 cell, and when it binds to and activates Sev, it induces the R7 precursor cell to differentiate into an R7 photoreceptor. The Sev protein is also expressed on several other precursor cells in the developing ommatidium, but none of these cells contact R8; therefore, the Sev protein is not activated, and these cells do not differentiate into R7 photoreceptors.



Figure 15–57 The assembly of photoreceptor cells in a developing *Drosophila* ommatidium. Cells are recruited sequentially to become photoreceptors, beginning with R8 and ending with R7, which is the last photoreceptor cell to develop.

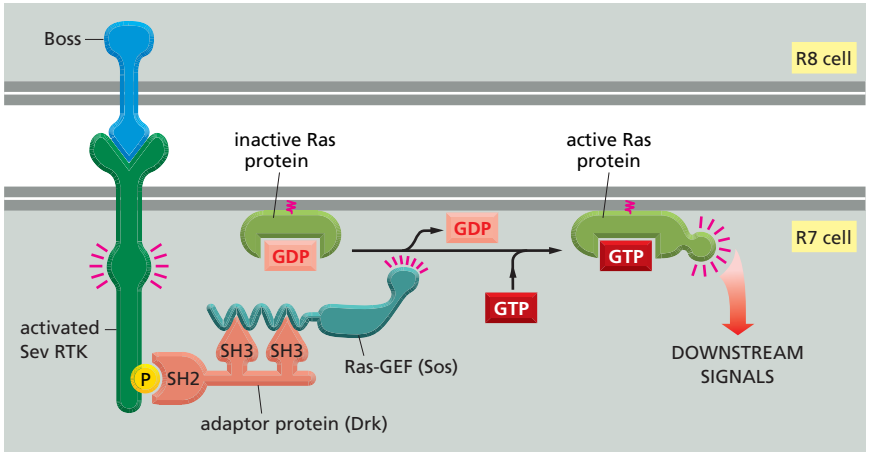


Figure 15–58 How the Sev RTK activates Ras in the fly eye. The activation of Sev on the surface of the R7 precursor cell by the Boss protein on the surface of R8 activates the Ras-GEF Sos via the adaptor protein Drk. Drk recognizes a specific phosphorylated tyrosine on the Sev protein by means of an SH2 domain and interacts with Sos by means of two SH3 domains. Sos stimulates the inactive Ras protein to replace its bound GDP by GTP, which activates Ras to relay the signal downstream, inducing the R7 precursor cell to differentiate into a UV-sensing photoreceptor cell.

The components of the intracellular signaling pathway activated by Sev in the R7 precursor cell proved more difficult to identify than the receptor or its ligand because mutations that inactivate them are lethal. The problem was solved by performing a genetic screen in flies with a partially inactive Sev protein. One of the genes identified encodes a Ras protein. Flies in which both copies of the *Ras* gene are inactivated by mutation die, whereas flies with only one inactivated copy survive. But when a fly carries a partially inactive Sev protein in its developing eyes, an inactivating mutation in one copy of the *Ras* gene results in the loss of R7. Moreover, if one of the *Ras* genes is rendered overactive by mutation, R7 develops even in mutants in which both *Sev* and *Boss* are inactive. These findings indicate that Ras acts downstream of Sev and that its activation in R7 precursor cells is necessary and sufficient to induce R7 differentiation.

A second gene identified in these genetic screens is called *Son-of-sevenless (Sos)*. It encodes a Ras-GEF, which is required for the Sev RTK to activate Ras. A third gene (called *Drk*) encodes an adaptor protein, which couples the Sev receptor to the Sos protein; the SH2 domain of the Drk adaptor binds to activated Sev, while its two SH3 domains bind to Sos (**Figure 15–58**). This type of genetic screen, in which organisms with a partially crippled component of a genetic pathway are used to identify genes encoding other proteins in the pathway, has now been widely used elsewhere with great success.

Biochemical and cell biological studies have shown that the coupling of RTKs to Ras occurs by a similar mechanism in mammalian cells, where the adaptor protein is called **Grb2** and the Ras-GEF is also called Sos (see Figure 15–22). Interestingly, when mammalian Sos activates Ras, Ras acts back to stimulate Sos further, forming a simple positive feedback loop.

RTKs are not the only means of activating Ras. Ca^{2+} and diacylglycerol, for example, activate a Ras-GEF that is found mainly in the brain; this Ras-GEF can couple GPCRs to Ras activation independently of Sos.

Once activated, Ras activates various other signaling proteins to relay the signal downstream, as we discuss next.

Ras Activates a MAP Kinase Signaling Module

Both the tyrosine phosphorylations and the activation of Ras triggered by activated RTKs are usually short-lived (**Figure 15–59**). *Tyrosine-specific protein phosphatases* quickly reverse the phosphorylations, and GAPs induce activated Ras to inactivate itself by hydrolyzing its bound GTP to GDP. To stimulate cells to proliferate or differentiate, these short-lived signaling events must be converted into longer-lasting ones that can sustain the signal and relay it downstream to the nucleus to alter the pattern of gene expression. One of the key mechanisms used for this purpose is a system of proteins called the *mitogen-activated protein kinase module* (**MAP kinase module**) (**Figure 15–60**). The three components of this system together form a functional signaling

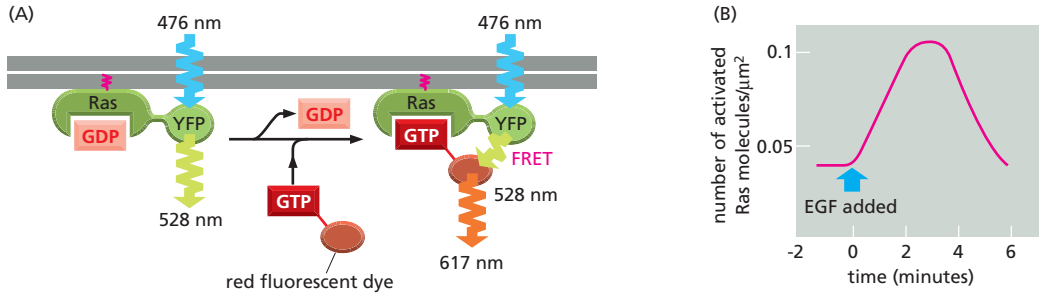


Figure 15–59 Transient activation of Ras revealed by single-molecule fluorescence resonance energy transfer (FRET). (A) Schematic drawing of the experimental strategy. Cells of a human cancer cell line are genetically engineered to express a Ras protein that is covalently linked to yellow fluorescent protein (YFP). GTP that is labeled with a red fluorescent dye is microinjected into some of the cells. The cells are then stimulated with the extracellular signal protein epidermal growth factor (EGF), and single fluorescent molecules of Ras-YFP at the inner surface of the plasma membrane are followed by video fluorescence microscopy in individual cells. When a fluorescent Ras-YFP molecule becomes activated, it exchanges unlabeled GDP for fluorescently labeled GTP; the yellow-green light emitted from the YFP now activates the fluorescent GTP to emit red light. Thus, the activation of single Ras molecules can be followed by the emission of red fluorescence from a previously yellow-green fluorescent spot at the plasma membrane. As shown in (B), activated Ras molecules can be detected after about 30 seconds of EGF stimulation. The red signal peaks at 3–4 minutes and then decreases to baseline by 6 minutes. As Ras-GAP is found to be recruited to the same spots at the plasma membrane as Ras, it presumably plays a major part in rapidly shutting off the Ras signal. (Modified from H. Murakoshi et al., *Proc. Natl Acad. Sci. U.S.A.* 101:7317–7322, 2004. With permission from National Academy of Sciences.)

module that has been highly conserved, from yeasts to humans, and is used, with variations, in many different signaling contexts.

The three components are all protein kinases. The final kinase in the series is called simply MAP kinase (MAPK). The next one upstream from this is MAP kinase kinase (MAPKK): it phosphorylates and thereby activates MAP kinase. And next above that, receiving an activating signal directly from Ras, is MAP kinase kinase kinase (MAPKKK): it phosphorylates and thereby activates MAPKK. In the mammalian **Ras-MAP-kinase signaling pathway**, these three kinases are known by shorter names: Raf (=MAPKKK), Mek (= MAPKK) and Erk (=MAPK).

Once activated, the MAP kinase relays the signal downstream by phosphorylating various proteins in the cell, including gene regulatory proteins and other protein kinases (see Figure 15–60). Erk MAP kinase, for example, enters the nucleus and phosphorylates one or more components of a gene regulatory complex. This activates the transcription of a set of *immediate early genes,* so named because they turn on within minutes after an RTK receives an extracellular signal,

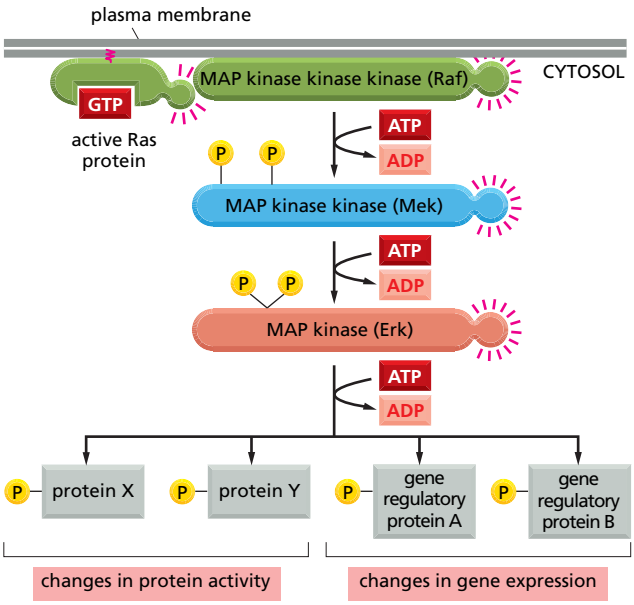


Figure 15–60 The MAP kinase serine/threonine phosphorylation module activated by Ras. The three-component module begins with a MAP kinase kinase kinase called *Raf.* Ras recruits Raf to the plasma membrane and helps activate it. Raf then activates the MAP kinase kinase *Mek,* which then activates the MAP kinase *Erk.* Erk in turn phosphorylates a variety of downstream proteins, including other protein kinases, as well as gene regulatory proteins in the nucleus. The resulting changes in gene expression and protein activity cause complex changes in cell behavior.

even if protein synthesis is experimentally blocked with drugs. Some of these genes encode other gene regulatory proteins that turn on other genes, a process that requires both protein synthesis and more time. (This relationship between the immediate early and later genes is similar to the relationship between primary and secondary response genes activated by the nuclear receptors discussed earlier—see Figure 15–15.)

In this way, the Ras–MAP-kinase signaling pathway conveys signals from the cell surface to the nucleus and alters the pattern of gene expression. Among the genes activated by this pathway are some that stimulate cell proliferation, such as the genes encoding *G_1 cyclins* (discussed in Chapter 17).

Extracellular signals usually activate MAP kinases only transiently, and the period during which the kinase remains active influences the response. When EGF activates its receptors in a neural precursor cell line, for example, Erk MAP kinase activity peaks at 5 minutes and rapidly declines, and the cells later go on to divide. By contrast, when NGF activates its receptors on the same cells, Erk activity remains high for many hours, and the cells stop proliferating and differentiate into neurons. Many factors can influence the duration of the signaling response, including positive and negative feedback loops (see Figure 15–28).

MAP kinases participate in both positive and negative feedback loops, which can combine to give responses that are either graded or switchlike and either brief or long lasting. In an example illustrated earlier, in Figure 15–24, MAP kinase activates a complex positive feedback loop to produce an all-or-none, irreversible response when frog oocytes are stimulated to mature by a brief exposure to the extracellular signal molecule progesterone. In many cells, MAP kinases activate a negative feedback loop by increasing the concentration of a *dual-specificity protein phosphatase* that removes the phosphate from both a tyrosine and a threonine to inactivate the MAP kinase (which is phosphorylated on the tyrosine and threonine by MAPKK). The increase in the phosphatase results from both an increase in the transcription of the phosphatase gene and the stabilization of the enzyme against degradation. In the Ras–MAP-kinase pathway shown in Figure 15–61, Erk also phosphorylates and inactivates Raf, providing another negative feedback loop that helps shut off the MAP kinase module.

Scaffold Proteins Help Prevent Cross-Talk Between Parallel MAP Kinase Modules

Three-component MAP kinase signaling modules operate in all eucaryotic cells, with different modules mediating different responses in the same cell. In budding yeast, for example, one such module mediates the response to mating pheromone, another the response to starvation, and yet another the response to osmotic shock. Some of these MAP kinase modules use one or more of the same kinases and yet manage to activate different effector proteins and hence different responses. As discussed earlier, one way in which cells avoid cross-talk between the different parallel signaling pathways and ensure that each response is specific is to use scaffold proteins (see Figure 15–21A). In yeast cells, such scaffolds bind all or some of the kinases in each MAP kinase module to form a complex and thereby help to ensure response specificity (**Figure 15–61**).

Mammalian cells also use this scaffold strategy to prevent cross-talk between different MAP kinase modules. At least 5 parallel MAP kinase modules can operate in a mammalian cell. These modules make use of at least 12 MAP kinases, 7 MAP kinase kinases, and 7 MAP kinase kinase kinases. Two of these modules (terminating in MAP kinases called JNK and p38) are activated by different kinds of cell stresses, such as UV irradiation, heat shock, and osmotic stress, as well as by inflammatory cytokines; others mainly mediate responses to signals from other cells.

Although the scaffold strategy provides precision and avoids cross-talk, it reduces the opportunities for amplification and spreading of the signal to different parts of the cell, which require at least some of the components to be diffusible (see Figure 15–17). It is unclear to what extent the individual components

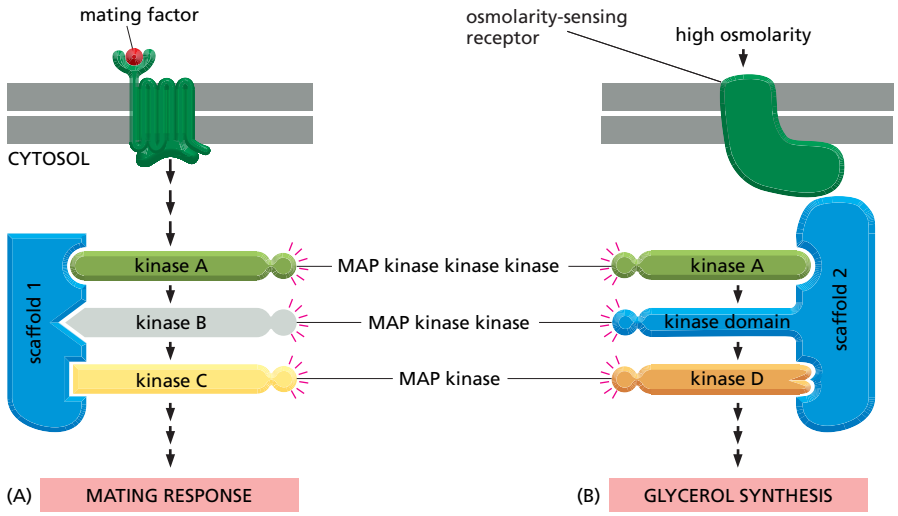


Figure 15–61 The organization of two MAP kinase modules by scaffold proteins in budding yeast. Budding yeast have at least six three-component MAP kinase modules involved in a variety of biological processes, including the two responses illustrated here—a mating response and the response to high osmolarity. (A) The mating response is triggered when a mating factor secreted by a yeast of opposite mating type binds to a GPCR. This activates a G protein, the βγ complex of which indirectly activates the MAP kinase kinase kinase (kinase A), which then relays the response onward. Once activated, the MAP kinase (kinase C) phosphorylates and thereby activates several proteins that mediate the mating response, in which the yeast cell stops dividing and prepares for fusion. The three kinases in this module are bound to scaffold protein 1. (B) In a second response, a yeast cell exposed to a high-osmolarity environment is induced to synthesize glycerol to increase its internal osmolarity. This response is mediated by a transmembrane, osmolarity-sensing, receptor protein and a different MAP kinase module bound to a second scaffold protein. (Note that the kinase domain of scaffold 2 provides the MAP kinase kinase activity of this module.) Although both pathways use the same MAP kinase kinase kinase (kinase A, *green*), there is no cross-talk between them, because the kinases in each module are bound to different scaffold proteins, and the osmosensor is bound to the same scaffold protein as the particular kinase it activates.

Remarkably, scaffold 1 is able to function even if it is stripped (by genetic engineering) of any ability to align or allosterically regulate its three bound kinases. Its role is apparently just to bring the kinases into close proximity, thereby increasing the rate at which they encounter each other, as expected for reaction rate accelerations caused by protein tethering (see Figure 3–80C).

of MAP kinase modules can dissociate from the scaffold during the activation process to permit amplification, which is why we refer to them as modules rather than cascades.

Rho Family GTPases Functionally Couple Cell-Surface Receptors to the Cytoskeleton

Besides the Ras proteins, the other class of Ras superfamily GTPases that relays signals from cell-surface receptors is the large **Rho family** (see Table 15–5). Rho family monomeric GTPases regulate both the actin and microtubule cytoskeletons, controlling cell shape, polarity, motility, and adhesion (discussed in Chapter 16); they also regulate cell-cycle progression, gene transcription, and membrane transport. They play a key part in the guidance of cell migration and nerve axon outgrowth, mediating cytoskeletal responses to the activation of a special class of guidance receptors. We focus on this aspect of Rho family function here.

The three best-characterized members are **Rho** itself, **Rac**, and **Cdc42**, each of which affects multiple downstream target proteins. In the same way as for Ras, GEFs activate and GAPs inactivate the Rho family GTPases; there are more than 60 Rho-GEFs and more than 70 Rho-GAPs in humans. Some of the GEFs and GAPs are specific for one particular family member, whereas others are less specific. Unlike Ras, which is membrane-associated even when inactive (with GDP bound), inactive Rho family GTPases are often bound to *guanine nucleotide dissociation inhibitors* (*GDIs*) in the cytosol, which prevent the GTPases from interacting with their Rho-GEFs at the plasma membrane.

Although cell-surface receptors activate the Rho GTPases by activating Rho-GEFs, it is not known in most cases how a receptor activates its GEF. One of the exceptions is an *Eph* RTK (see Figure 15–52) on the surface of motor neurons that helps guide the migrating tip of the axon (called a *growth cone*) to its muscle target. The binding of a cell-surface *ephrin* protein activates the Eph receptor, causing the growth cones to collapse, thereby repelling them from inappropriate regions and keeping them on track. The response depends on a Rho-GEF called *ephexin*, which is stably associated with the cytosolic tail of the Eph receptor. When ephrin binding activates the Eph receptor, the receptor activates a cytoplasmic tyrosine kinase that phosphorylates ephexin on a tyrosine, enhancing the ability of ephexin to activate the Rho protein RhoA. The activated RhoA (RhoA-GTP) then regulates various downstream target proteins, including some effector proteins that control the actin cytoskeleton, causing the growth cone to collapse (**Figure 15–62**).

Ephrins are among the best-characterized extracellular *guidance proteins*, with many functions, both in the nervous system and outside it. One of their roles, for example, is to direct the way in which nerve axons grow from the eye to the optic tectum so as to create a neural 'map' of the visual field in the brain

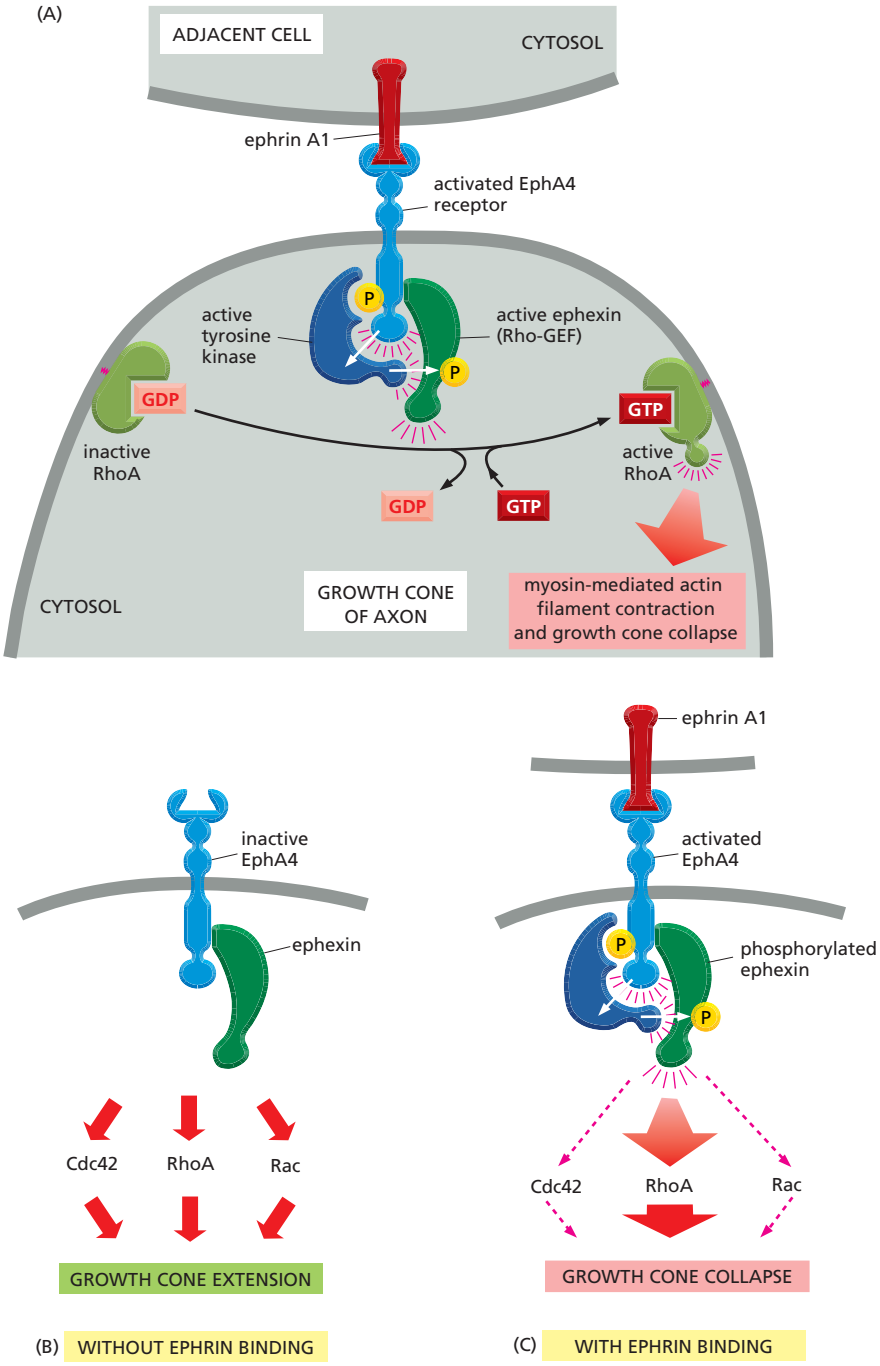


Figure 15–62 Growth cone collapse mediated by Rho family GTPases. (A) The binding of transmembrane ephrin A1 proteins on an adjacent cell activates EphA4 RTKs on the growth cone of an axon, by the mechanism illustrated in Figure 15–53A. Phosphotyrosines on the activated Eph receptors recruit and activate a cytoplasmic tyrosine kinase to phosphorylate the receptor-associated Rho-GEF ephexin on a tyrosine. This enhances the ability of the ephexin to activate the Rho family GTPase RhoA. RhoA then induces the growth cone to collapse by stimulating the myosin-dependent contraction of the actin cytoskeleton. (B) When ephrin A1 is not bound to the EphA4 receptor, ephexin activates three different Rho family members (Cdc42, Rac, and RhoA) about equally, promoting the forward advance of the growth cone. (C) EphrinA1 binding to EphA4 alters the activity of ephexin so that it now mainly activates RhoA, causing the growth cone to collapse.

(discussed in Chapter 22). But the guidance of axon outgrowth is a complex matter, and other types of guidance receptors are also involved, as discussed in Chapter 22. All of these receptors, however, are thought to guide cell movement by influencing the cytoskeleton, and to do so via members of the Rho family.

Having considered how RTKs use GEFs and monomeric GTPases to relay signals into the cell, we now consider a second major strategy that RTKs use that depends on a quite different intracellular relay mechanism.

PI 3-Kinase Produces Lipid Docking Sites in the Plasma Membrane

As mentioned earlier, one of the proteins that binds to the intracellular tail of RTK molecules is the plasma-membrane-bound enzyme **phosphoinositide 3-kinase (PI 3-kinase)**. This kinase principally phosphorylates inositol phospho-

lipids rather than proteins, and both RTKs and GPCRs can activate it. It plays a central part in promoting cell survival and growth.

Phosphatidylinositol (PI) is unique among membrane lipids because it can undergo reversible phosphorylation at multiple sites on its inositol head group to generate a variety of phosphorylated PI lipids called **phosphoinositides** (see Figure 15–37). When activated, PI 3-kinase catalyzes phosphorylation at the 3 position of the inositol ring to generate several phosphoinositides (**Figure 15–63**). The production of PI(3,4,5)P_3 matters most because it can serve as a docking site for various intracellular signaling proteins, which assemble into signaling complexes that relay the signal into the cell from the cytosolic face of the plasma membrane (see Figure 15–21C).

Notice the difference between this use of phosphoinositides and their use described earlier, in which PI(4,5)P_2 is cleaved by PLCβ (in the case of GPCRs) or PLCγ (in the case of RTKs) to generate soluble IP_3 and membrane-bound diacylglycerol (see Figures 15–38 and 15–39). By contrast, PI(3,4,5)P_3 is not cleaved by PLC. It is made from PI(4,5)P_2 and then remains in the plasma membrane until specific *phosphoinositide phosphatases* dephosphorylate it. Prominent among these is the *PTEN* phosphatase, which dephosphorylates the 3 position of the inositol ring. Mutations in PTEN are found in many cancers: by prolonging signaling by PI 3-kinase, they promote uncontrolled cell growth.

There are various types of PI 3-kinases. Those activated by RTKs and GPCRs belong to class I. These are heterodimers composed of a common catalytic subunit and different regulatory subunits. RTKs activate *class Ia PI 3-kinases*, in which the regulatory subunit is an adaptor protein that binds to two phosphotyrosines on activated RTKs through its two SH2 domains (see Figure 15–55A). GPCRs activate *class Ib PI 3-kinases*, which have a regulatory subunit that binds to the βγ complex of an activated trimeric G protein when GPCRs are activated by their extracellular ligand. The direct binding of activated Ras can also activate the common class I catalytic subunit.

Intracellular signaling proteins bind to PI(3,4,5)P_3 produced by activated PI 3-kinase via a specific interaction domain, such as a **pleckstrin homology (PH) domain** mentioned earlier, first identified in the platelet protein pleckstrin. PH domains function mainly as protein–protein interaction domains, and it is only a small subset of them that bind to PIP_3; at least some of these also recognize a specific membrane-bound protein as well as the PIP_3, which greatly increases the specificity of the binding and helps to explain why the signaling proteins with PIP_3-binding PH domains do not all dock at all PI(3,4,5)P_3 sites. PH domains occur in about 200 human proteins, including the Ras-GEF Sos discussed earlier (see Figure 15–22).

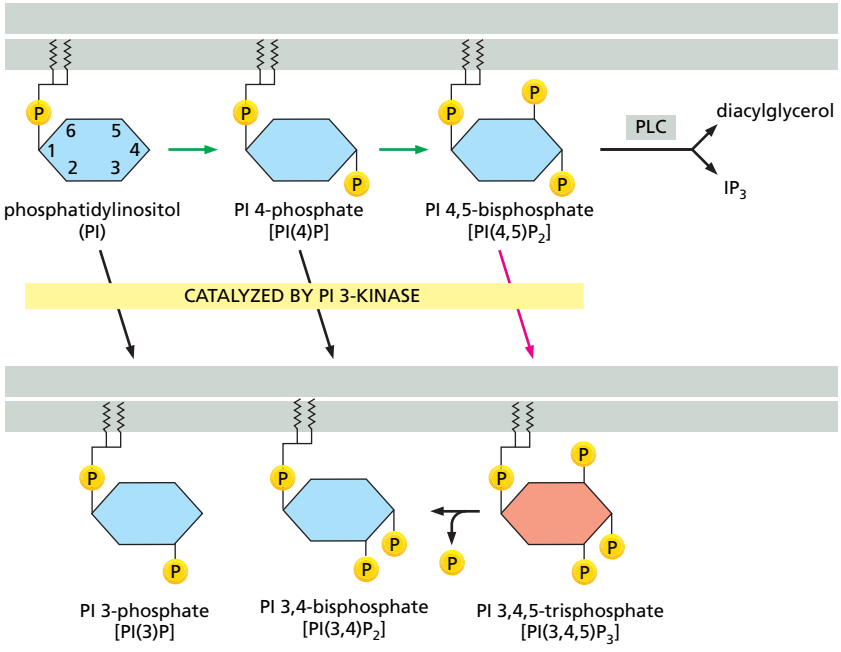


Figure 15–63 The generation of phosphoinositide docking sites by PI 3-kinase. PI 3-kinase phosphorylates the inositol ring on carbon atom 3 to generate the phosphoinositides shown at the bottom of the figure (diverting them away from the pathway leading to IP_3 and diacylglycerol). The most important phosphorylation (indicated in *red*) is of PI(4,5)P_2 to PI(3,4,5)P_3, which can serve as a docking site for signaling proteins with PIP_3-binding PH domains. Other inositol phospholipid kinases catalyze the phosphorylations indicated by the *green arrows*.

One especially important PH-domain-containing protein is the serine/threonine protein kinase *Akt*. The *PI-3-kinase–Akt signaling pathway* is the major pathway activated by the hormone *insulin*. It also plays a key part in promoting the survival and growth of many cell types in both invertebrates and vertebrates, as we now discuss.

The PI-3-Kinase–Akt Signaling Pathway Stimulates Animal Cells to Survive and Grow

For a multicellular organism to grow, its cells must grow (accumulate mass and enlarge). If the cells simply divided without growing, they would get progressively smaller and the organism as a whole would stay the same size. As discussed earlier, extracellular signals are usually required for animal cells to grow and divide, as well as to survive (see Figure 15–8). Members of the *insulin-like growth factor (IGF)* family of signal proteins, for example, stimulate many types of animal cells to survive and grow. They bind to specific RTKs (see Figure 15–52), which activate PI 3-kinase to produce $PI(3,4,5)P_3$. The PIP_3 recruits two protein kinases to the plasma membrane via their PH domains—**Akt** (also called *protein kinase B*, or *PKB)* and *phosphoinositide-dependent protein kinase 1* (*PDK1*), and this leads to the activation of Akt (**Figure 15–64**). Once activated, Akt phosphorylates various target proteins at the plasma membrane, as well as in the cytosol and nucleus. The effect on most of the known targets is to inactivate them; but the targets are such that these actions of Akt all conspire to enhance cell survival and growth.

One effect of Akt, for example, is to phosphorylate a cytosolic protein called *Bad,* which, in its nonphosphorylated state, promotes cell death by apoptosis (discussed in Chapter 18). The phosphorylation of Bad by Akt creates phosphoserine-binding sites for a scaffold protein called *14-3-3*, which sequesters phosphorylated Bad and keeps it out of action, thereby promoting cell survival (see Figure 15–64).

The **PI-3-kinase–Akt pathway** signals cells to grow through a more complex mechanism that depends on a large, serine/threonine protein kinase called **TOR** (named as the target of *rapamycin*, a bacterial toxin that inactivates the kinase and is used clinically as both an immunosuppressant and anti-cancer drug). TOR was originally identified in yeasts in genetic screens for rapamycin resistance; in mammalian cells, it is called **mTOR**. TOR exists in cells in two functionally distinct multiprotein complexes. In mammalian cells, *mTOR complex 1* contains the protein *raptor*; this complex is sensitive to rapamycin, and it stimulates cell

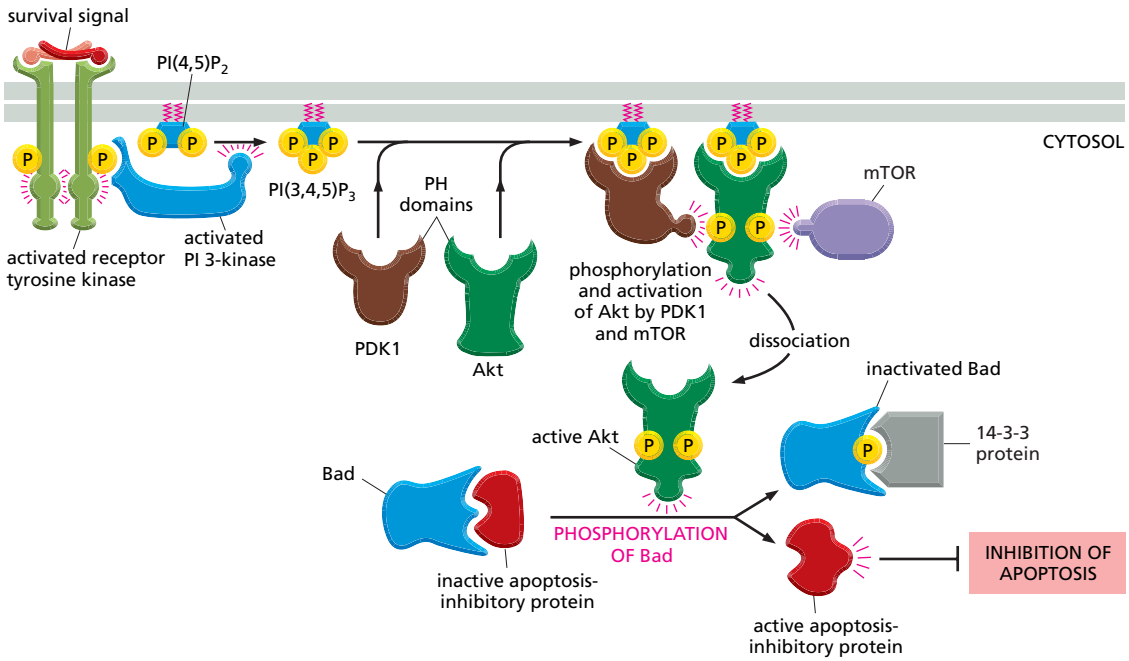


Figure 15–64 One way in which signaling through PI 3-kinase promotes cell survival. An extracellular survival signal activates an RTK, which recruits and activates PI 3-kinase. The PI 3-kinase produces $PI(3,4,5)P_3$, which serves as a docking site for two serine/threonine kinases with PH domains—Akt and the phosphoinositol-dependent kinase PDK1—which are brought into proximity at the plasma membrane. The Akt is phosphorylated on a serine by a third kinase (usually mTOR), which alters the conformation of the Akt so that it can be phosphorylated on a threonine by PDK1, which activates the Akt. The activated Akt now dissociates from the plasma membrane and phosphorylates various target proteins, including the Bad protein.

When unphosphorylated, Bad holds one or more apoptosis-inhibitory proteins (of the Bcl2 family—discussed in Chapter 18) in an inactive state. Once phosphorylated, Bad releases the inhibitory proteins, which now can block apoptosis and thereby promote cell survival. As shown, once phosphorylated, Bad binds to a ubiquitous cytosolic protein called *14-3-3,* which keeps Bad out of action.

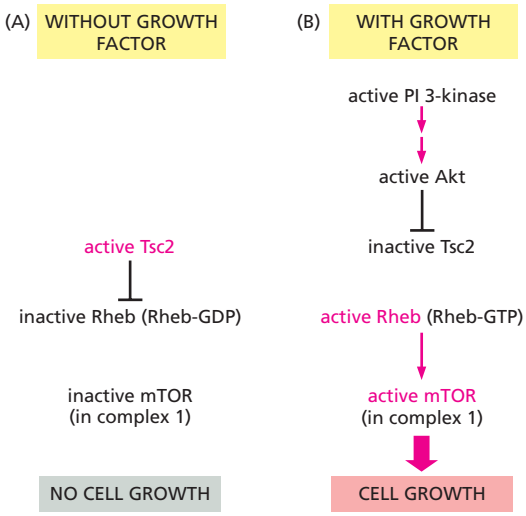


Figure 15–65 Activation of mTOR by the PI-3-kinase–Akt signaling pathway. (A) In the absence of extracellular growth factors, Tsc2 (a Rheb-GAP) keeps Rheb inactive; mTOR in complex I is inactive, and there is no cell growth. (B) In the presence of growth factors, activated Akt phosphorylates and inhibits Tsc2, thereby promoting the activation of Rheb. Activated Rheb (Rheb-GTP) helps activate mTOR in complex 1, which in turn stimulates cell growth. Figure 15–64 shows how growth factors (or survival signals) activate Akt.

The Erk MAP kinase (see Figure 15–60) can also phosphorylate and inhibit Tsc2 and thereby activate mTOR. Thus, both the PI-3-kinase–Akt and Ras–MAP-kinase signaling pathways converge on mTOR in complex 1 to stimulate cell growth.

Tsc2 is short for *tuberous sclerosis protein 2,* and it is one component of a heterodimer composed of Tsc1 and Tsc2 (not shown); these proteins are so called because mutations in either gene encoding them cause the genetic disease *tuberous sclerosis,* which is associated with benign tumors in the brain and elsewhere that contain abnormally large cells.

growth—both by promoting ribosome production and protein synthesis and by inhibiting protein degradation. Complex 1 also promotes both cell growth and cell survival by stimulating nutrient uptake and metabolism. *mTOR complex 2* contains the protein *rictor* and is insensitive to rapamycin; it helps to activate Akt (see Figure 15–64), and it regulates the actin cytoskeleton via Rho family GTPases.

The mTOR in complex 1 integrates inputs from various sources, including extracellular signal proteins referred to as *growth factors* and nutrients such as amino acids, both of which help activate mTOR and promote cell growth. The growth factors activate mTOR mainly via the PI-3-kinase–Akt pathway. Akt activates mTOR in complex 1 indirectly by phosphorylating, and thereby inhibiting, a GAP called Tsc2. Tsc2 acts on a monomeric Ras-related GTPase called **Rheb** (see Table 15–5, p. 926). Rheb in its active form (Rheb-GTP) activates mTOR. The net result is that Akt activates mTOR and thereby promotes cell growth (**Figure 15–65**). We discuss how mTOR stimulates ribosome production and protein synthesis in Chapter 17 (see Figure 17–65).

The Downstream Signaling Pathways Activated By RTKs and GPCRs Overlap

As mentioned earlier, RTKs and GPCRs activate some of the same intracellular signaling pathways. Both, for example, can activate the inositol phospholipid pathway triggered by phospholipase C. Moreover, even when they activate different pathways, the different pathways can converge on the same target proteins. **Figure 15–66** illustrates both of these types of signaling overlaps: it summarizes five parallel intracellular signaling pathways that we have discussed so far—one triggered by GPCRs, two triggered by RTKs, and two triggered by both kinds of receptors.

Tyrosine-Kinase-Associated Receptors Depend on Cytoplasmic Tyrosine Kinases

Many cell-surface receptors depend on tyrosine phosphorylation for their activity and yet lack a tyrosine kinase domain. These receptors act through **cytoplasmic tyrosine kinases**, which are associated with the receptors and phosphorylate various target proteins, often including the receptors themselves, when the receptors bind their ligand. These **tyrosine-kinase-associated receptors** thus function in much the same way as RTKs, except that their kinase domain is encoded by a separate gene and is noncovalently associated with the receptor polypeptide chain. A variety of receptor classes belong in this category, including the receptors for antigen and interleukins on lymphocytes (discussed in Chapter

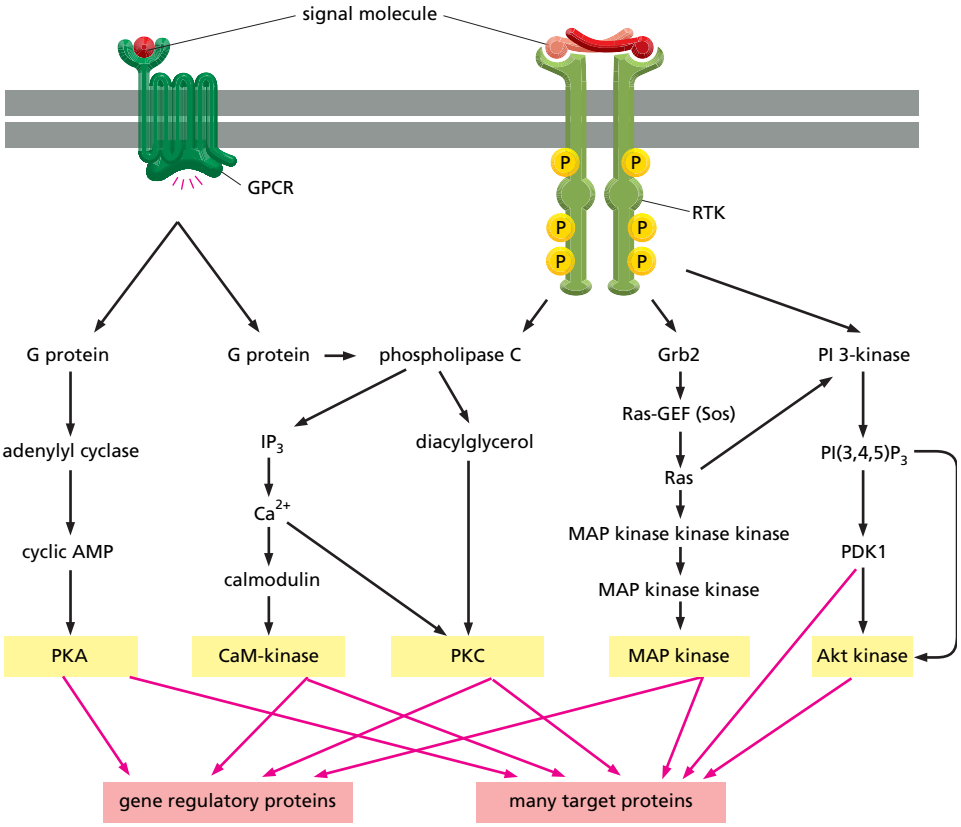


Figure 15–66 Five parallel intracellular signaling pathways activated by GPCRs, RTKs, or both. In this hypothetical example, the five kinases (shaded *yellow*) at the end of each signaling pathway phosphorylate target proteins (shaded *red*), many of which are phosphorylated by more than one of the kinases. The phospholipase C activated by the two types of receptors is different: GPCRs activate PLCβ, whereas RTKs activate PLCγ (not shown). Although not shown, some GPCRs can also activate Ras, but they do so independently of Grb2, via a Ras-GEF that is activated by Ca^{2+} and diacylglycerol.

25), integrins (discussed in Chapter 19), and receptors for various cytokines and some hormones. As with RTKs, many of these receptors are either preformed dimers (**Figure 15–67**) or are cross-linked into dimers by ligand binding.

Some of these receptors depend on members of the largest family of mammalian cytoplasmic tyrosine kinases, the **Src family** (see Figures 3–10 and 3–69), which includes *Src, Yes, Fgr, Fyn, Lck, Lyn, Hck,* and *Blk*. These protein kinases all contain SH2 and SH3 domains and are located on the cytoplasmic side of the plasma membrane, held there partly by their interaction with transmembrane receptor proteins and partly by covalently attached lipid chains. Different family members are associated with different receptors and phosphorylate overlapping but distinct sets of target proteins. Lyn, Fyn, and Lck, for example, are each associated with different sets of receptors in lymphocytes (discussed in Chapter 25). In each case, the kinase is activated when an extracellular ligand binds to the appropriate receptor protein. Src itself, as well as several other family members, can also bind to activated RTKs; in these cases, the receptor and cytoplasmic kinases mutually stimulate each other's catalytic activity, thereby strengthening and prolonging the signal (see Figure 15–62). There are even some G proteins (G_s and G_i) that can activate Src, which is one way that the activation of GPCRs can lead to tyrosine phosphorylation of intracellular signaling proteins and effector proteins.

Another type of cytoplasmic tyrosine kinase associates with *integrins,* the main receptors that cells use to bind to the extracellular matrix (discussed in Chapter 19). The binding of matrix components to integrins activates intracellu-

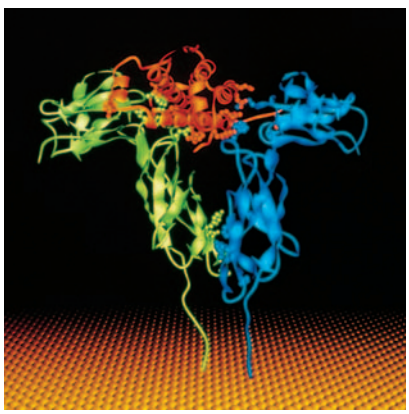


Figure 15–67 The three-dimensional structure of human growth hormone bound to its receptor. <CGCT> The receptor is a homodimer, and the hormone *(red)* binds to each of the two identical subunits (one shown in *green* and the other in *blue*), which recognize different parts of the monomeric hormone. Hormone binding is thought to cause a rearrangement of the subunits, which activates cytoplasmic tyrosine kinases that are tightly bound to the cytosolic tails of the receptors (not shown). The structure shown was determined by x-ray crystallographic studies of complexes formed between the hormone and extracellular receptor domains produced by recombinant DNA technology. (From A.M. deVos, M. Ultsch and A.A. Kossiakoff, *Science* 255:306–312, 1992. With permission from AAAS.)

lar signaling pathways that influence the behavior of the cell. When integrins cluster at sites of matrix contact, they help trigger the assembly of cell–matrix junctions called *focal adhesions*. Among the many proteins recruited into these junctions is the cytoplasmic tyrosine kinase called **focal adhesion kinase (FAK)**, which binds to the cytosolic tail of one of the integrin subunits with the assistance of other proteins. The clustered FAK molecules cross-phosphorylate each other, creating phosphotyrosine-docking sites where the Src kinase can bind. Src and FAK now phosphorylate each other and other proteins that assemble in the junction, including many of the signaling proteins used by RTKs. In this way, the two tyrosine kinases signal to the cell that it has adhered to a suitable substratum, where the cell can now survive, grow, divide, migrate, and so on.

The largest and most diverse class of receptors that rely on cytoplasmic tyrosine kinases to relay signals into the cell is the class of *cytokine receptors*, which we consider next.

Cytokine Receptors Activate the JAK–STAT Signaling Pathway, Providing a Fast Track to the Nucleus

The large family of **cytokine receptors** includes receptors for many kinds of local mediators (collectively called *cytokines*), as well as receptors for some hormones, such as *growth hormone* (see Figure 15–67) and *prolactin*. These receptors are stably associated with cytoplasmic tyrosine kinases called **Janus kinases (JAKs)** (after the two-faced Roman god), which phosphorylate and activate gene regulatory proteins called **STATs** (**s**ignal **t**ransducers and **a**ctivators of **t**ranscription). STAT proteins are located in the cytosol and are referred to as *latent gene regulatory proteins* because they only migrate into the nucleus and regulate gene transcription after they are activated.

Although many intracellular signaling pathways lead from cell-surface receptors to the nucleus, where they alter gene transcription (see Figure 15–66), the **JAK–STAT signaling pathway** provides one of the more direct routes. Cytokine receptors are dimers or trimers and are stably associated with one or two of the four known JAKs (JAK1, JAK2, JAK3, and Tyk2). Cytokine binding alters the arrangement so as to bring two JAKs into close proximity so that they transphosphorylate each other, thereby increasing the activity of their tyrosine kinase domains. The JAKs then phosphorylate tyrosines on the cytokine receptors, creating phosphotyrosine docking sites for STATs (**Figure 15–68**). Some adaptor proteins can also bind to some of these sites and couple cytokine receptors to the Ras–MAP-kinase signaling pathway discussed earlier, but these will not be discussed here.

There are at least six STATs in mammals. Each has an SH2 domain that performs two functions. First, it mediates the binding of the STAT protein to a phosphotyrosine docking site on an activated cytokine receptor. Once bound, the JAKs phosphorylate the STAT on tyrosines, causing the STAT to dissociate from the receptor. Second, the SH2 domain on the released STAT now mediates its binding to a phosphotyrosine on another STAT molecule, forming either a STAT homodimer or a heterodimer. The STAT dimer then translocates to the nucleus, where, in combination with other gene regulatory proteins, it binds to a specific DNA response element in various genes and stimulates their transcription (see Figure 15–68). In response to the hormone prolactin, for example, which stimulates breast cells to produce milk, activated STAT5 stimulates the transcription of genes that encode milk proteins. **Table 15–6** lists some of the more than 30 cytokines and hormones that activate the JAK–STAT pathway by binding to cytokine receptors; it also shows the specific JAKs and STATs involved.

Some STATs also have an SH2 domain that enables them to dock onto specific phosphotyrosines on some activated RTKs. These receptors can directly activate the bound STAT, independently of JAKs. The nematode *C. elegans* uses STATs for signaling but does not make any JAKs or cytokine receptors, suggesting that STATs evolved before JAKs and cytokine receptors.

Negative feedback regulates the responses mediated by the JAK–STAT pathway. In addition to activating genes that encode proteins mediating the

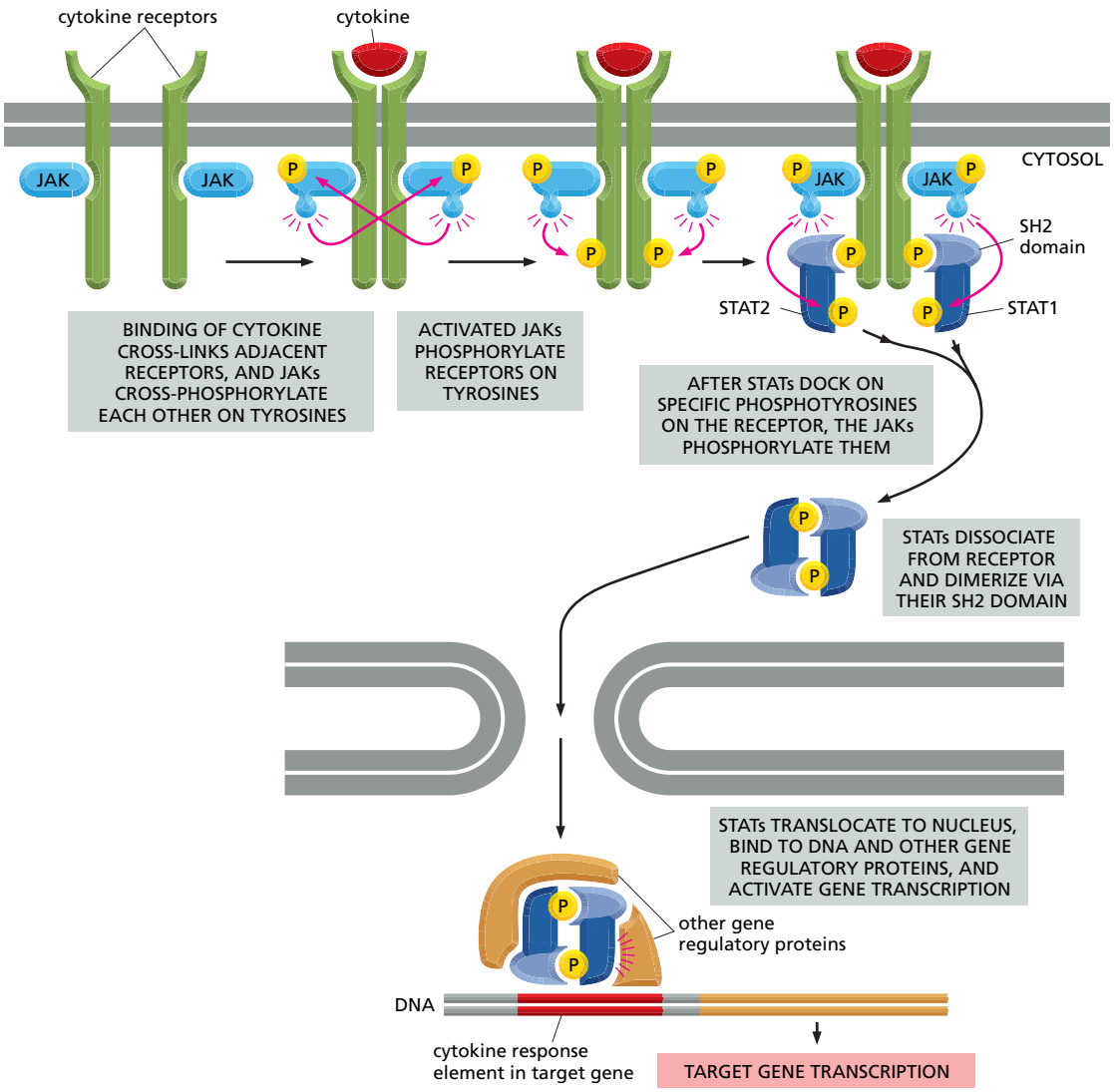


Figure 15–68 The JAK–STAT signaling pathway activated by cytokines. The binding of the cytokine either causes two separate receptor polypeptide chains to dimerize (as shown) or reorients the receptor chains in a preformed dimer. In either case, the associated JAKs are brought together so that they can cross-phosphorylate each other on tyrosines, starting the signaling process shown. In some cases, the active receptor is a trimer rather than a dimer.

cytokine-induced response, the STAT dimers can also activate genes that encode inhibitory proteins that help shut off the response. Some of these proteins bind to and inactivate phosphorylated JAKs and their associated phosphorylated receptors; others bind to phosphorylated STAT dimers and prevent them from binding to their DNA targets. Such negative feedback mechanisms, however, are not enough on their own to turn off the response. Inactivation of the activated JAKs and STATs requires dephosphorylation of their phosphotyrosines.

Protein Tyrosine Phosphatases Reverse Tyrosine Phosphorylations

In all signaling pathways that use tyrosine phosphorylation, the tyrosine phosphorylations are reversed by dephosphorylation performed by **protein tyrosine phosphatases**. These phosphatases are as important in the signaling process as the protein tyrosine kinases that add the phosphates. Whereas only a few types of *serine/threonine protein phosphatase* catalytic subunits are responsible for removing phosphate groups from phosphorylated serines and threonines on proteins, there are about 100 protein tyrosine phosphatases encoded in the human genome, including some *dual-specificity phosphatases* that also dephosphorylate serines and threonines.

Like tyrosine kinases, the tyrosine phosphatases occur in both cytoplasmic and transmembrane forms, none of which are structurally related to serine/threonine protein phosphatases. Some of the transmembrane forms are thought to act as cell-surface receptors, but as this has not been established, they are generally referred to as receptorlike tyrosine phosphatases.

Table 15–6 Some Extracellular Signal Proteins That Act Through Cytokine Receptors and the JAK–STAT Signaling Pathway

SIGNAL PROTEIN	RECEPTOR-ASSOCIATED JAKs	STATS ACTIVATED	SOME RESPONSES
γ-interferon	JAK1 and JAK2	STAT1	activates macrophages
α-interferon	Tyk2 and JAK2	STAT1 and STAT2	increases cell resistance to viral infection
Erythropoietin	JAK2	STAT5	stimulates production of erythrocytes
Prolactin	JAK1 and JAK2	STAT5	stimulates milk production
Growth hormone	JAK2	STAT1 and STAT5	stimulates growth by inducing IGF1 production
GMCSF	JAK2	STAT5	stimulates production of granulocytes and macrophages

Unlike serine/threonine protein phosphatases, which generally have broad specificity, most tyrosine phosphatases display exquisite specificity for their substrates, removing phosphate groups from only selected phosphotyrosines on a subset of proteins. Together, these phosphatases ensure that tyrosine phosphorylations are short-lived and that the level of tyrosine phosphorylation in resting cells is very low. They do not, however, simply continuously reverse the effects of protein tyrosine kinases; they are often regulated to act only at the appropriate time and place in a signaling response or in the cell-division cycle (discussed in Chapter 17).

Having discussed the crucial role of tyrosine phosphorylation and dephosphorylation in the intracellular signaling pathways activated by many enzyme-coupled receptors, we now turn to a class of enzyme-coupled receptors that rely entirely on serine/threonine phosphorylation. These *receptor serine/ threonine kinases* activate an even more direct signaling pathway to the nucleus than does the JAK–STAT pathway. They directly phosphorylate latent gene regulatory proteins called *Smads,* which then translocate into the nucleus to activate gene transcription.

Signal Proteins of the TGFβ Superfamily Act Through Receptor Serine/Threonine Kinases and Smads

The **transforming growth factor-β (TGFβ) superfamily** consists of a large number (30–40 in humans) of structurally related, secreted, dimeric proteins. They act either as hormones or, more commonly, as local mediators to regulate a wide range of biological functions in all animals. During development, they regulate pattern formation and influence various cell behaviors, including proliferation, specification and differentiation, extracellular matrix production, and cell death. In adults, they are involved in tissue repair and in immune regulation, as well as in many other processes. The superfamily consists of the TGFβ/*activin* family and the larger *bone morphogenetic protein (BMP)* family.

All of these proteins act through enzyme-coupled receptors that are single-pass transmembrane proteins with a serine/threonine kinase domain on the cytosolic side of the plasma membrane. There are two classes of these **receptor serine/threonine kinases**—*type I* and *type II*—which are structurally similar homodimers. Each member of the TGFβ superfamily binds to a characteristic combination of type-I and type-II receptor dimers, bringing the kinase domains together so that the type-II receptor can phosphorylate and activate the type-I receptor, forming an active tetrameric receptor complex.

Once activated, the receptor complex uses a strategy for rapidly relaying the signal to the nucleus that is very similar to the JAK–STAT strategy used by cytokine receptors. The activated type-I receptor directly binds and phosphorylates a latent gene regulatory protein of the **Smad family** (named after the first two identified, Sma in *C. elegans* and Mad in *Drosophila*). Activated TGFβ/activin receptors phosphorylate Smad2 or Smad3, while activated BMP receptors phosphorylate Smad1, Smad5, or Smad8. Once one of these *receptor-activated Smads (R-Smads)* has been phosphorylated, it dissociates from the receptor and binds to Smad4 (called a *co-Smad),* which can form a complex

with any of the five R-Smads. The Smad complex then translocates into the nucleus, where it associates with other gene regulatory proteins and regulates the transcription of specific target genes (**Figure 15–69**). Because the partner proteins in the nucleus vary depending on the cell type and state of the cell, the genes affected vary.

Activated TGFβ receptors and their bound ligand are endocytosed by two distinct routes, one leading to further activation, and the other leading to inactivation. The activation route depends on clathrin-coated vesicles and leads to early endosomes (discussed in Chapter 13), where most of the Smad activation occurs. An anchoring protein called *SARA* (for *Smad anchor for receptor activation*) has an important role in this pathway; it is concentrated in early endosomes and binds to both activated TGFβ receptors and Smads, increasing the efficiency of receptor-mediated Smad phosphorylation. The inactivation route depends on *caveolae* (discussed in Chapter 13) and leads to receptor ubiquitylation and degradation in proteasomes.

Some TGFβ family members serve as graded morphogens during development, inducing different responses in a developing cell depending on the concentration of the morphogen (see Figure 15–10, and discussed in Chapter 22). Their effective extracellular concentrations are often regulated by secreted inhibitory proteins that bind directly to the signal molecules and prevent them from activating their receptors on target cells. *Noggin* and *chordin,* for example, inhibit BMPs, and *follistatin* inhibits activins. Some of these inhibitors, as well as most TGFβ family members themselves, are secreted as inactive precursors, which are activated by extracellular proteolytic cleavage.

During the signaling response, the Smads continuously shuttle between the cytoplasm and the nucleus: they are dephosphorylated in the nucleus and exported to the cytoplasm, where they can be rephosphorylated by activated receptors. In this way, the effect exerted on the target genes reflects both the concentration of the extracellular signal and the time for which it continues to act on the cell-surface receptors (often several hours). Cells exposed to a morphogen at

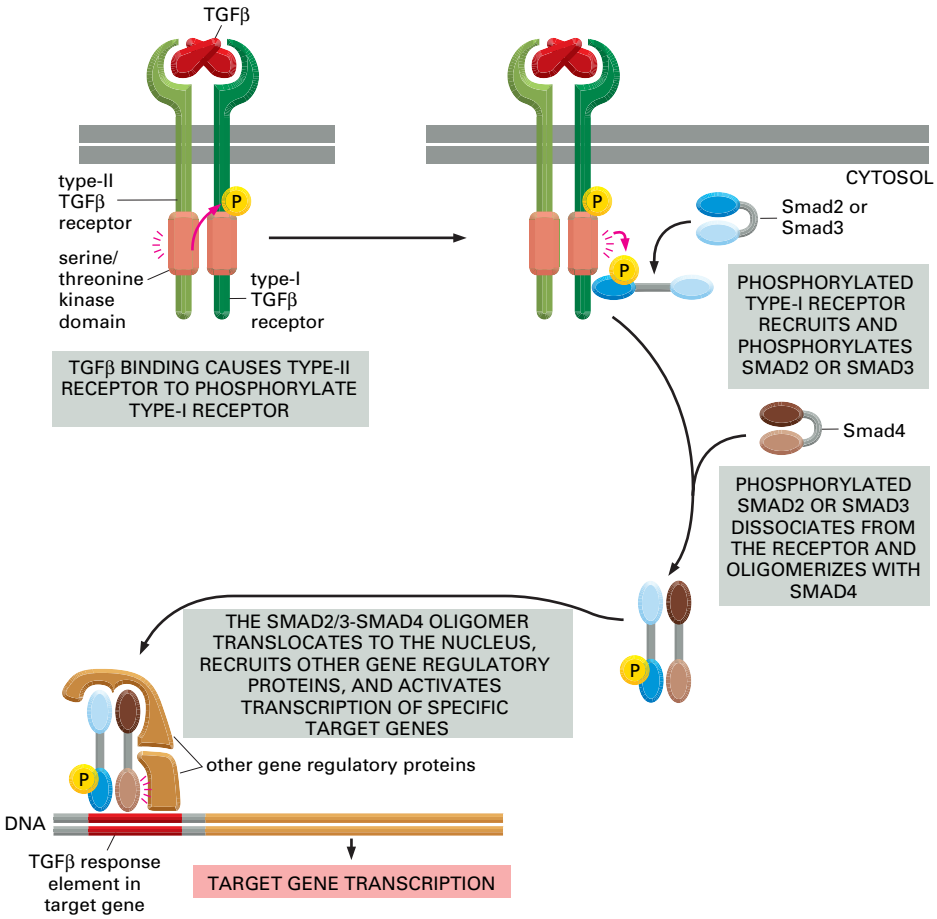


Figure 15–69 The Smad-dependent signaling pathway activated by TGFβ. Note that TGFβ is a dimer and that Smads open up to expose a dimerization surface when they are phosphorylated. Several features of the pathway have been omitted for simplicity, including the following. (1) The type-I and type-II receptor proteins are both thought to be homodimers. (2) The type-I receptors are normally associated with an inhibitory protein, which dissociates when the type-I receptor is phosphorylated by a type-II receptor. (3) The individual Smads are thought to be trimers. (4) An anchoring protein called SARA helps to recruit Smad2 or Smad3 to the activated type I receptor, mainly in endosomes, as discussed in the text.

high concentration, or for a long time, or both, will switch on one set of genes, while cells receiving a lower or more transient exposure will switch on another set.

As with the JAK–STAT pathway, negative feedback regulates the Smad pathway. Among the target genes activated by Smad complexes are those that encode *inhibitory Smads,* either Smad6 or Smad7. Smad7 (and possibly Smad6) binds to the activated receptor and inhibits its signaling ability in at least three ways: (1) it competes with R-Smads for binding sites on the receptor, decreasing R-Smad phosphorylation; (2) it recruits a ubiquitin ligase called *Smurf,* which ubiquitylates the receptor, leading to receptor internalization and degradation (it is because Smurfs also ubiquitylate and promote the degradation of Smads that they are called *Smad ubiquitylation regulatory factors,* or Smurfs); and (3) it recruits a protein phosphatase that dephosphorylates and inactivates the receptor. In addition, the inhibitory Smads bind to the co-Smad, Smad4, and inhibit it, either by preventing its binding to R-Smads or by promoting its ubiquitylation and degradation.

Although receptor serine/threonine kinases operate mainly through the Smad pathway just described, they can also affect other intracellular signaling pathways. Conversely, signaling proteins in other pathways can phosphorylate Smads and thereby influence signaling along the Smad pathway.

Serine/Threonine and Tyrosine Protein Kinases Are Structurally Related

All the signaling pathways activated by GPCRs and enzyme-coupled receptors that we have discussed so far depend on serine/threonine-specific protein kinases, tyrosine-specific protein kinases, or both. These kinases are all structurally related, as summarized in **Figure 15–70**.

The bewildering complexity of multiple cross-regulatory signaling pathways and feedback loops that we have discussed is not just a haphazard tangle, but a highly evolved system for processing and interpreting the mass of signals that impinge upon animal cells. The whole molecular control network, leading from the receptors at the cell surface to the genes in the nucleus, can be viewed as a computing device; and, like that other computing device, the brain, it presents one of the hardest problems in biology. We can identify the components and discover how they work individually. We can understand how small subsets of components work together as regulatory modules, as we have seen. But, it is a much more difficult task to understand how the system works as a whole. This is not only because the system is complex; it is also because the way it behaves is strongly dependent on the quantitative details of the molecular interactions, and for most animal cells we have only rough qualitative information.

In bacterial cells, the signaling pathways are simpler, and precise quantitative information is much easier to obtain. It is therefore possible to give a detailed account of how a complete signaling system works, at least for one particular bacterial cell behavior and the signals that control it. We discuss here one such example, in which bacteria respond to environmental signals delivered via enzyme-coupled receptors that are again kinases, but of a type unrelated to those we have discussed so far.

Bacterial Chemotaxis Depends on a Two-Component Signaling Pathway Activated by Histidine-Kinase-Associated Receptors

As pointed out earlier, many of the mechanisms involved in chemical signaling between cells in multicellular animals are thought to have evolved from mechanisms used by unicellular organisms to respond to chemical changes in their environment. In fact, both types of organisms use some of the same intracellular mediators, such as cyclic nucleotides and Ca^{2+}. Among the best-studied reactions of unicellular organisms to extracellular signals are their *chemotactic responses,* in which cell movement is oriented toward or away

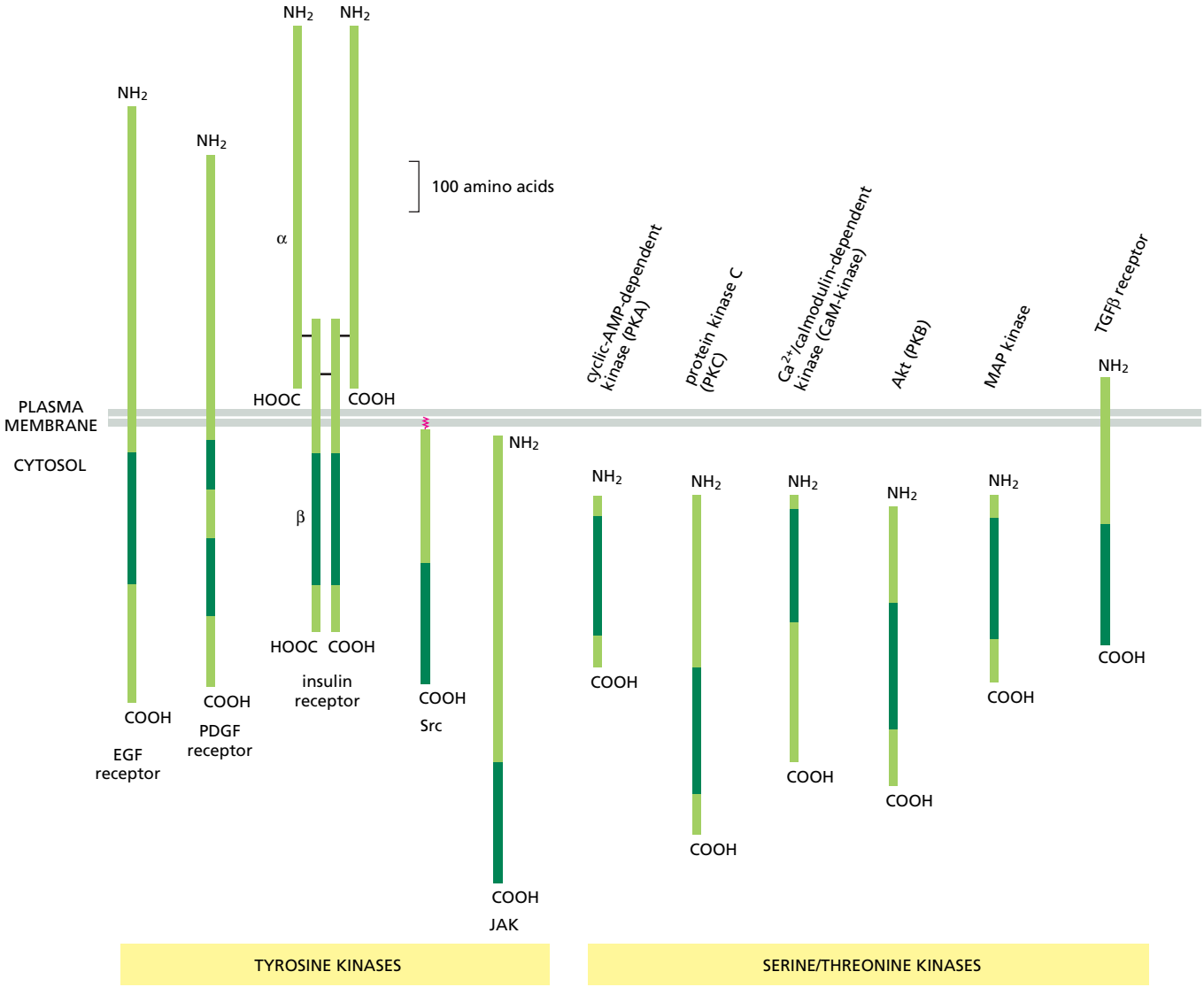


Figure 15–70 Some of the protein kinases discussed in this chapter. The size and location of their catalytic domains *(dark green)* are shown. In each case the catalytic domain is about 250 amino acids long. These domains are all similar in amino acid sequence, suggesting that they have all evolved from a common primordial kinase (see also Figure 3–66). Note that all of the tyrosine kinases shown are bound to the plasma membrane (JAKs are bound by their association with cytokine receptors—see Figure 15–68), whereas most of the serine/threonine kinases are in the cytosol.

from a source of some chemical in the environment. We conclude this section on enzyme-coupled receptors with a brief account of **bacterial chemotaxis**, which provides a particularly well-understood illustration of the crucial role of adaptation in the response to chemical signals. This chemotaxis response is mediated by **histidine-kinase-associated receptors** that activate a *two-component signaling pathway*, which is also used by yeasts and plants, although apparently not by animals.

Motile bacteria such as *E. coli* will swim toward higher concentrations of nutrients *(attractants)*, including sugars, amino acids, and small peptides, and away from higher concentrations of various noxious chemicals *(repellents)*. They swim by means of four to six flagella, each of which is attached by a short, flexible hook at its base to a small protein disc embedded in the bacterial membrane. This disc is part of a tiny motor that uses the energy stored in the transmembrane H^+ gradient to rotate rapidly and turn the helical flagellum (**Figure 15–71**). Because the flagella on the bacterial surface have an intrinsic "handedness," different directions of rotation have different effects on movement. The flagella spend most of the time rotating counterclockwise, which draws all the flagella together into a coherent bundle, so that the bacterium swims uniformly in one direction. In the absence of any environmental stimulus, every second or so, one or more of the motors transiently reverses direction so that the attached flagellum breaks out of the bundle, causing the bacterium to tumble chaotically without moving forward (**Figure 15–72**). This sequence produces a characteristic pattern of movement in which smooth swimming in a straight line is interrupted by abrupt, random changes in direction caused by tumbling.

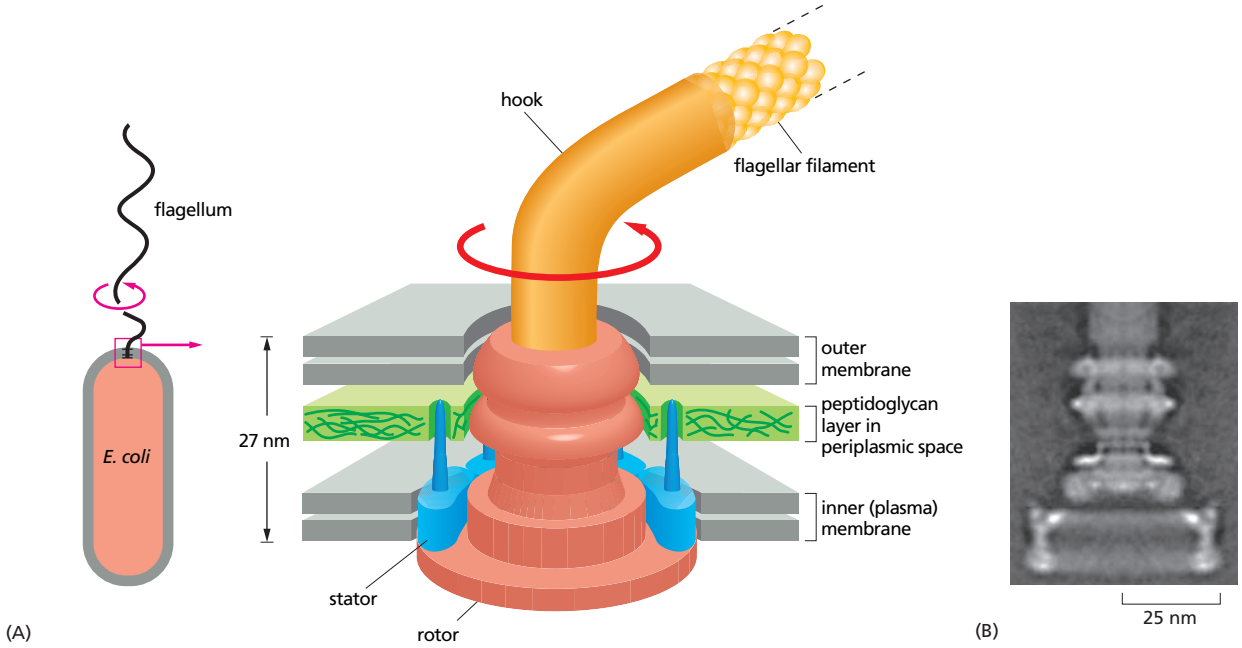


Figure 15–71 The bacterial flagellar motor. (A) A schematic drawing of the structure. The flagellum is linked to a flexible hook. The hook is attached to a series of protein rings (shown in *orange*), which are embedded in the outer and inner (plasma) membranes. The rings form a rotor, which rotates with the flagellum at more than 100 revolutions per second. The rotation is driven by a flow of protons through an outer ring of proteins, the stator *(blue)*, which is embedded in the inner membrane and tethered to the peptidoglycan layer. The stator also contains the proteins responsible for switching the direction of rotation. (B) A flagellar motor reconstructed from multiple electron microscopic images. (A, based on data from T. Kubori et al., *J. Mol. Biol.* 226:433–446,1992, with permission from Academic Press, and N.R. Francis et al., *Proc. Natl Acad. Sci. U.S.A.* 89:6304–6308, 1992, with permission from National Academy of Sciences; B, from D. Thomas, D.G. Morgan and D.J. DeRosier, *J. Bacteriol.* 183:6404–6412, 2001. With permission from American Society for Microbiology.)

The normal swimming behavior of bacteria is modified by chemotactic attractants or repellents, which bind to specific receptor proteins and affect the frequency of tumbling by increasing or decreasing the time that elapses between successive changes in direction of flagellar rotation. When bacteria are swimming in a favorable direction (toward a higher concentration of an attractant or away from a higher concentration of a repellent), they tumble less frequently than when they are swimming in an unfavorable direction (or when no gradient is present). Since the periods of smooth swimming are longer when a bacterium is traveling in a favorable direction, it will gradually progress in that direction—toward an attractant or away from a repellent.

Histidine-kinase-associated **chemotaxis receptors** mediate these responses. The receptors typically are dimeric transmembrane proteins that bind specific attractants and repellents on the outside of the plasma membrane. The cytoplasmic tail of the receptor is stably associated with a histidine kinase *CheA* via an adaptor protein *CheW* (**Figure 15–73**). Repellent binding activates the receptors, whereas attractant binding inactivates them; a single receptor can bind either type of molecule, with opposite consequences. The binding of a repellent to the receptor activates *CheA*, which phosphorylates itself on a histidine and almost immediately transfers the phosphoryl group to an aspartic acid on a response regulator protein *CheY*. The phosphorylated *CheY* dissociates from the receptor, diffuses through the cytosol, binds to a flagellar motor, and causes the motor to rotate clockwise, so that the bacterium tumbles. CheY has intrinsic phosphatase activity and dephosphorylates itself in a process that is greatly accelerated by the *CheZ* protein (see Figure 15–73).

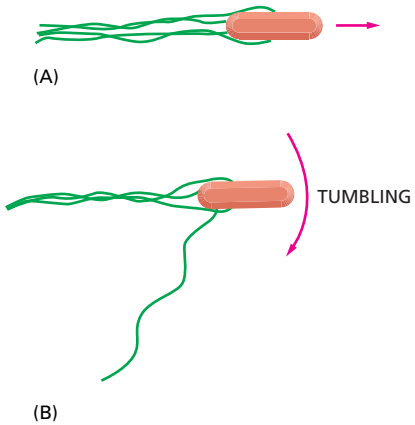


Figure 15–72 Positions of the flagella on *E. coli* during swimming. (A) When the flagella rotate counterclockwise (as seen looking into the cell from the flagellum), they are drawn together into a single bundle, which acts as a propeller to produce smooth swimming. (B) When one or more motors reverses direction, the attached flagellum breaks out of the bundle, causing the bacterium to tumble.

Receptor Methylation Is Responsible for Adaptation in Bacterial Chemotaxis

The change in tumbling frequency in response to an increase in the concentration of an attractant or repellent occurs within less than a second, but it is only

transient. Even if the higher level of ligand is maintained, within minutes, the bacteria *adapt* (desensitize), to the increased stimulus. The adaptation is a crucial part of the response, as it enables a bacterium to compare its present environment with that in its recent past and thereby respond to *changes* in the concentration of ligand rather than to steady-state levels.

The adaptation is mediated by the covalent methylation (catalyzed by a *methyl transferase*) or demethylation (catalyzed by a *methylase*) of the chemotaxis receptors, which change their responsiveness to ligand binding as a consequence of the covalent modification. When an attractant binds to a chemotaxis receptor, for example, it has two effects: (1) It decreases the ability of the receptor to activate CheA, resulting in a decreased rate of tumbling. (2) It slowly (over minutes) alters the receptor so that it can be methylated by the methyl transferase, which returns the receptor's ability to activate CheA to its original level. Thus, the unmethylated receptor without a bound ligand has the same activity as the methylated receptor with a bound ligand, and the tumbling frequency of the bacterium is therefore the same in both cases.

Each receptor dimer has eight methylation sites, and the number of sites methylated increases with increasing concentration of the attractant (as each receptor spends a longer time with a ligand bound at higher concentrations). When the attractant is removed, the methylase demethylates the receptor. Although the level of receptor methylation changes during a chemotactic response, it remains constant once a bacterium is adapted because an exact balance is reached between the rates of methylation and demethylation. A simple model of how ligand binding and methylation may operate in bacterial chemotaxis proposes that both receptor methylation and repellent binding tighten the structure of the multisubunit receptor and its associated signaling proteins, thereby increasing signaling and tumbling; by contrast, receptor demethylation and attractant binding loosen the structure of the complex, thereby decreasing signaling and tumbling. It is thought that the sensitivity of this response is greatly increased by cooperative effects that result from the clustering of the cytoplasmic tails of adjacent receptors in the membrane (**Figure 15–74**).

All of the genes and proteins involved in bacterial chemotaxis have now been identified, and their interactions have been measured in some detail. It therefore seems likely that it will be the first signaling system to be completely understood in molecular terms. Even in this relatively simple signaling network, however, computer-based simulations are required to comprehend how the system works as an integrated network. Cell signaling pathways will provide an especially rich area of investigation for a new generation of computational biologists, as the network properties of these pathways are not understandable without powerful computational tools.

As mentioned earlier, there are some cell-surface receptor proteins that do not fit into the three major classes we have discussed thus far—ion-channel-coupled, G-protein-coupled, and enzyme-coupled. In the next section, we consider cell-surface receptors that activate signaling pathways that depend on proteolysis to regulate the activity of latent gene regulatory proteins. These pathways have especially important roles in animal development and in tissue renewal and repair.

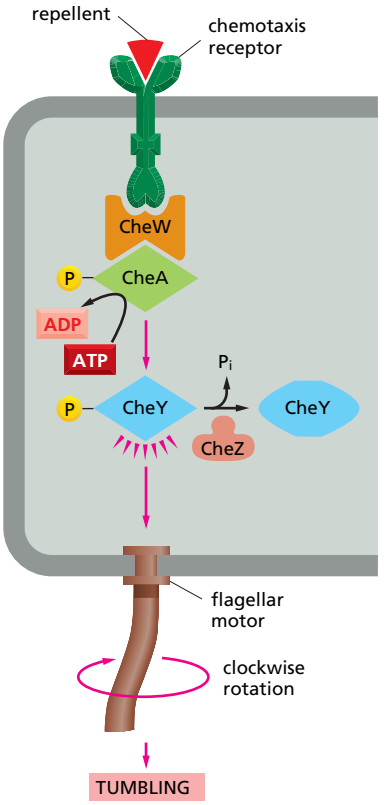
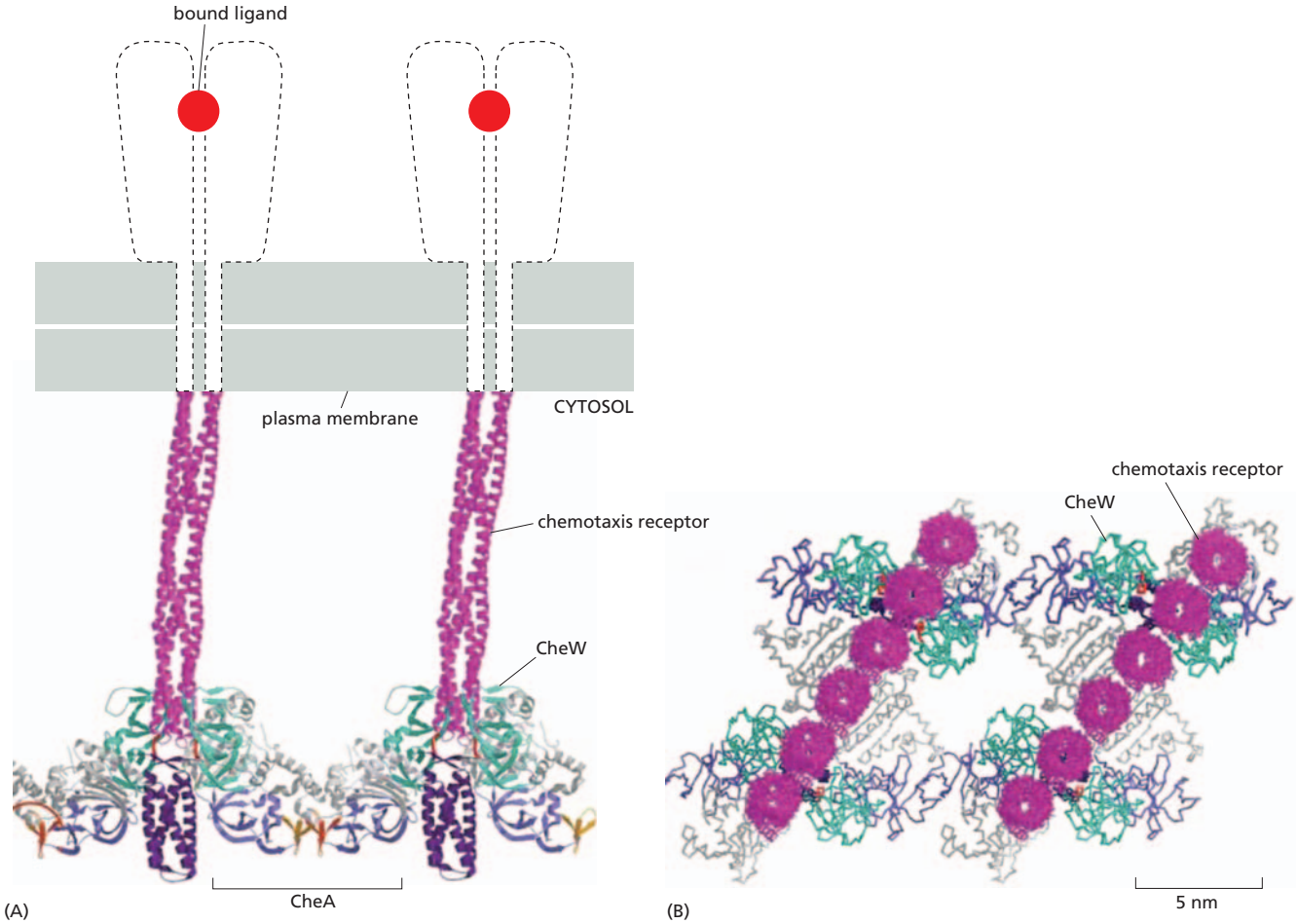


Figure 15–73 The two-component signaling pathway that enables chemotaxis receptors to control the flagellar motors during bacterial chemotaxis. The histidine kinase CheA is stably bound to the receptor via the adaptor protein CheW. The receptors and their associated proteins are all clustered at one end of the cell (see Figure 15–74). The binding of a repellent increases the activity of the receptor, which stimulates CheA to phosphorylate itself on histidine. CheA quickly transfers its covalently bound, high-energy phosphoryl group directly to the response regulator CheY to generate CheY-phosphate, which then diffuses away, binds to a flagellar motor, and causes the motor to rotate clockwise, which results in tumbling. The binding of an attractant has the opposite effect: it decreases the activity of the receptor and therefore decreases the phosphorylation of CheA and CheY, which results in counterclockwise flagellar rotation and smooth swimming. CheZ accelerates the autodephosphorylation of CheY-phosphate, thereby inactivating CheY. Each of the phosphorylated intermediates decays with a half life of less than a second, enabling the bacterium to respond very quickly to changes in its environment (see Figure 15–11).

Summary

There are various classes of enzyme-coupled receptors, including receptor tyrosine kinases (RTKs), tyrosine-kinase-associated receptors, receptor serine/threonine kinases, and histidine-kinase-associated receptors. The first two classes are by far the most numerous in mammals.

Ligand binding to RTKs induces the receptors to cross-phosphorylate their cytoplasmic domains on multiple tyrosines. This transautophosphorylation both stimulates the kinases and produces a set of phosphotyrosines that serve as docking sites for a set of intracellular signaling proteins, which bind via their SH2 (or PTB) domains. One such signaling protein serves as an adaptor to couple some activated receptors to a Ras-GEF (Sos), which activates the monomeric GTPase Ras; Ras, in turn, activates a



three-component MAP kinase signaling module, which relays the signal to the nucleus by phosphorylating gene regulatory proteins there. Another important signaling protein that can dock on activated RTKs is PI 3-kinase, which phosphorylates specific phosphoinositides to produce lipid docking sites in the plasma membrane for signaling proteins with phosphoinositide-binding PH domains, including the serine/threonine protein kinase Akt (PKB), which plays a key part in the control of cell survival and growth. Many receptor classes, including some RTKs, activate Rho family monomeric GTPases, which functionally couple the receptors to the cytoskeleton.

Tyrosine-kinase-associated receptors depend on various cytoplasmic tyrosine kinases for their action. These kinases include members of the Src family, which associate with many kinds of receptors, and the focal adhesion kinase (FAK), which associates with integrins at focal adhesions. The cytoplasmic tyrosine kinases then phosphorylate a variety of signaling proteins to relay the signal onward. The largest family of receptors in this class is the cytokine receptor family. When stimulated by ligand binding, these receptors activate JAK cytoplasmic tyrosine kinases, which phosphorylate STATs. The STATs then dimerize, translocate to the nucleus, and activate the transcription of specific genes. Receptor serine/threonine kinases, which are activated by signal proteins of the TGFβ superfamily, act similarly: they directly phosphorylate and activate Smads, which then oligomerize with another Smad, translocate to the nucleus, and activate gene transcription.

Bacterial chemotaxis is regulated by a signaling pathway that is understood exceptionally well. It is governed by histidine-kinase-associated chemotaxis receptors, which activate a two-component signaling pathway. When activated by a repellent, a chemotaxis receptor stimulates its associated histidine kinase to phosphorylate itself on histidine and then transfer the phosphoryl group to a response regulator protein, which relays the signal to a flagellar motor to alter the bacterium's swimming behavior. Attractants have the opposite effect on this kinase and therefore on swimming. These responses depend crucially on receptor adaptation, which is mediated by reversible receptor methylation.

Figure 15–74 A structural model for the clustering of chemotaxis receptors in the bacterial plasma membrane. Two clustered receptors are shown. Each is a homodimer, and the long α-helical cytoplasmic tail of each of the two subunits folds back on itself. Thus, each subunit contributes two helices to the four-helix bundle. Distance measurements between protein domains, determined by a technique known as pulsed electron-spin resonance, have been integrated with the known three-dimensional protein structure to produce the model shown. The clustering of receptor-CheA-CheW complexes into arrays of this type allows cooperative interactions between adjacent complexes, greatly increasing the sensitivity of the signaling process. (A) View of the structure just below the bacterial plasma membrane. The extracellular domain of the receptor is drawn schematically. (B) View of the same structure looking inward from the plasma membrane. (Adapted from S.Y. Park et al., *Nat. Struct. Biol.* 13:400–407, 2006. With permission from Macmillan Publishers Ltd.)

SIGNALING PATHWAYS DEPENDENT ON REGULATED PROTEOLYSIS OF LATENT GENE REGULATORY PROTEINS

The need for intercellular signaling is never greater than during animal development. Each cell in the embryo has to be guided along one developmental pathway or another according to its history, its position, and the character of its neighbors. At each step in the pathway, it must exchange signals with its neighbors to coordinate its behavior with theirs, ensuring the correct number and pattern of different cell types in each tissue and organ. Most of the signaling pathways already discussed are widely used for these developmental purposes, controlling cell survival, growth, proliferation, adhesion, specification, differentiation, and migration.

There are other signaling pathways, however, that are at least as important in controlling developmental processes but relay signals in other ways from cell-surface receptors to the interior of the cell. Several of these pathways depend on *regulated proteolysis* to control the activity and location of *latent gene regulatory proteins*, which enter the nucleus and activate the transcription of specific target genes only after they have been signaled to do so. Although the STAT and Smad proteins discussed earlier are also latent gene regulatory proteins, they are activated by phosphorylation in response to extracellular signals rather than by highly selective protein degradation. Signaling pathways that use latent gene regulatory proteins provide, as their primary function, a relatively direct linear pathway by which extracellular signals can control gene expression, which is presumably why they are so commonly used during development, especially to control cell fate decisions.

Although most of the pathways we discuss in this section were discovered through genetic studies in *Drosophila,* they have been highly conserved in evolution and are used over and over again during the development of different tissues and different animals. As we discuss in Chapter 23, they also have a crucial role in the many developmental processes that continue in adult tissues and organs, where new cells are constantly being produced.

We discuss four of these signaling pathways in this section: that mediated by the receptor protein *Notch,* that activated by secreted *Wnt* proteins, that activated by secreted *Hedgehog* proteins, and the pathway that activates the latent gene regulatory protein *NFκB*. All of these pathways have crucial roles in animal development, and we discuss the central roles of Notch, Wnt, and Hedgehog signaling in embryonic development in detail in Chapter 22.

The Receptor Protein Notch Is a Latent Gene Regulatory Protein

Signaling through the **Notch** receptor protein may be the most widely used signaling pathway in animal development. As discussed in Chapter 22, it has a general role in controlling cell fate choices and regulating pattern formation during the development of most tissues, as well as in the continual renewal of tissues such as the lining of the gut. It is best known, however, for its role in the production of nerve cells in *Drosophila*, which usually arise as isolated single cells within an epithelial sheet of precursor cells. During this process, when a precursor cell commits to becoming a nerve cell, it signals to its immediate neighbors not to do the same; the inhibited cells develop into epidermal cells instead. This process, called *lateral inhibition*, depends on a contact-dependent signaling mechanism that is activated by a single-pass transmembrane signal protein called **Delta**, displayed on the surface of the future neural cell. By binding to the Notch receptor protein on a neighboring cell, Delta signals to the neighbor not to become neural (**Figure 15–75**). When this signaling process is defective, the neighbors of neural cells also develop as neural cells, producing a huge excess of neurons at the expense of epidermal cells, which is lethal.

Signaling between adjacent cells via Notch and Delta (or Deltalike ligands) regulates cell fate choices in many tissues and animals. Often, it mediates lateral

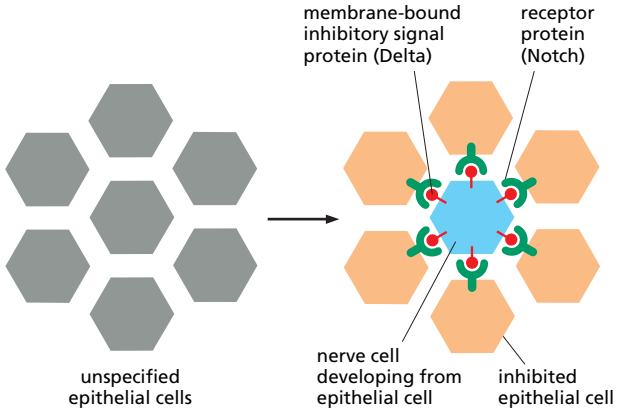


Figure 15–75 Lateral inhibition mediated by Notch and Delta during nerve cell development in *Drosophila*. When individual cells in the epithelium begin to develop as neural cells, they signal to their neighbors not to do the same. This inhibitory, contact-dependent signaling is mediated by the ligand Delta, which appears on the surface of the future nerve cell and binds to Notch receptor proteins on the neighboring cells. In many tissues, all the cells in a cluster initially express both Delta and Notch, and a competition occurs, with one cell emerging as winner, expressing Delta strongly and inhibiting its neighbors from doing likewise (see Figure 22–60). In other cases, additional factors interact with Delta or Notch to make some cells susceptible to the lateral inhibition signal and others unresponsive to it.

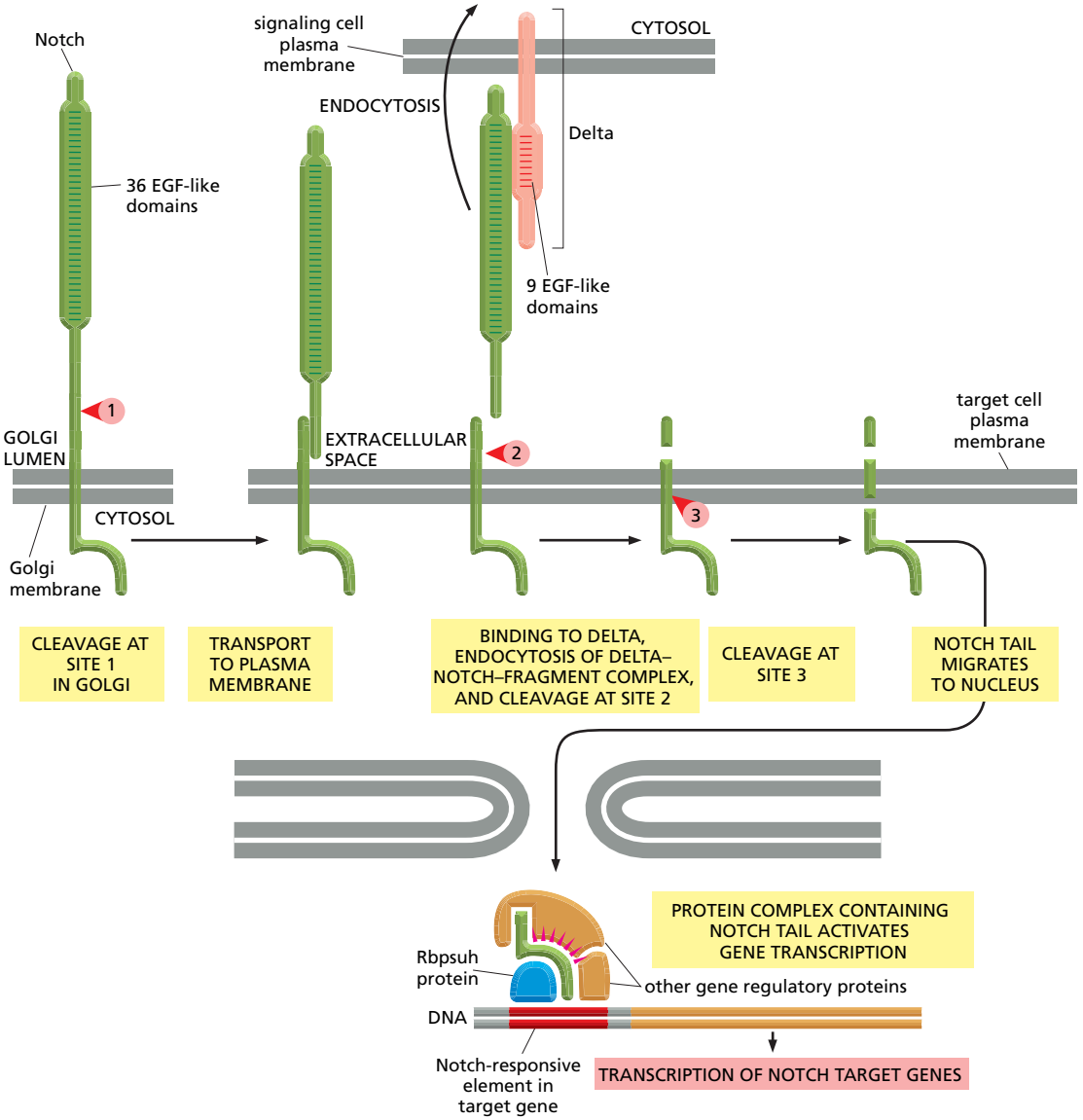
inhibition to control the formation of mixtures of differentiated cell types within a tissue, as in the fly nervous system. In some other cases, however, it works in the opposite way, promoting rather than inhibiting a particular cell fate and driving neighboring cells to behave similarly. There is hardly any developmental cell behavior that is not regulated by Notch signaling in one tissue or another.

Notch is a single-pass transmembrane protein that requires proteolytic processing to function. It acts as a latent gene regulatory protein and provides the simplest and most direct signaling pathway known from a cell-surface receptor to the nucleus. When activated by the binding of Delta on another cell, a plasma-membrane-bound protease cleaves off the cytoplasmic tail of Notch, and the released tail translocates into the nucleus to activate the transcription of a set of Notch-response genes. The Notch tail fragment acts by binding to a DNA-binding protein, converting it from a transcriptional repressor into a transcriptional activator. We shall see that the Wnt and Hedgehog signaling pathways use this same strategy of switching a transcriptional repressor into a transcriptional activator to regulate cell fate. The set of genes regulated by Notch signaling varies depending on the tissue and circumstances, although the primary targets in most cells are members of a gene family, known (in mammals) as *Hes* genes, which encode inhibitory gene regulatory proteins. In the nervous system, for example, the products of the *Hes* genes block the expression of genes required for neural differentiation.

The Notch receptor undergoes three successive proteolytic cleavage steps, but only the last two depend on Delta binding. As part of its normal biosynthesis, it is cleaved in the Golgi apparatus to form a heterodimer, which is then transported to the cell surface as the mature receptor. The binding of Delta to Notch induces a second cleavage in the extracellular domain, mediated by an extracellular protease. A final cleavage quickly follows, cutting free the cytoplasmic tail of the activated receptor (**Figure 15–76**). Note that, unlike most receptors, the activation of Notch is irreversible; once activated by ligand binding, the protein cannot be used again.

This final cleavage of the Notch tail occurs just within the transmembrane segment, and it is mediated by a protease complex called *γ-secretase*, which is also responsible for the intramembrane cleavage of various other proteins. One of its essential subunits is *Presenilin*, so-called because mutations in the gene encoding it are a frequent cause of early-onset, familial Alzheimer's disease, a form of presenile dementia. The protease complex is thought to contribute to this and other forms of Alzheimer's disease by generating extracellular peptide fragments from a transmembrane neuronal protein; the fragments accumulate in excessive amounts and form aggregates of misfolded protein called amyloid plaques, which may injure nerve cells and contribute to their degeneration and loss.

Both Notch and Delta are glycoproteins, and their interaction is regulated by the glycosylation of Notch. The *Fringe family* of glycosyltransferases, in particular, adds extra sugars to the *O*-linked oligosaccharide (discussed in Chapter 13) on Notch, which alters the specificity of Notch for its ligands. This has provided the first example of the modulation of ligand-receptor signaling by differential receptor glycosylation.



Wnt Proteins Bind to Frizzled Receptors and Inhibit the Degradation of β-Catenin

Wnt proteins are secreted signal molecules that act as local mediators and morphogens to control many aspects of development in all animals that have been studied. They were discovered independently in flies and in mice: in *Drosophila*, the *Wingless (Wg)* gene originally came to light because of its role as a morphogen in wing development (discussed in Chapter 22), while in mice, the *Int1* gene was found because it promoted the development of breast tumors when activated by the integration of a virus next to it. Both of these genes encode Wnt proteins. Wnts are unusual as secreted proteins in that they have a fatty acid chain covalently attached to their N-terminus, which increases their binding to cell surfaces. There are 19 Wnts in humans, each having distinct, but often overlapping, functions.

Wnts can activate at least three types of intracellular signaling pathways: (1) The *Wnt/β-catenin pathway* (also known as the *canonical Wnt pathway*) is centered on the latent gene regulatory protein *β-catenin*. (2) The *planar polarity pathway* coordinates the polarization of cells in the plane of a developing epithelium (discussed in Chapters 19 and 22) and depends on Rho family GTPases. (3) The Wnt/Ca^{2+} pathway stimulates an increase of intracellular Ca^{2+}, with consequences of the sort we described earlier for other pathways. All three pathways begin with the binding of Wnts to **Frizzled** family cell-surface receptors, which

Figure 15–76 The processing and activation of Notch by proteolytic cleavage. The *numbered red arrowheads* indicate the sites of proteolytic cleavage. The first proteolytic processing step occurs within the *trans* Golgi network to generate the mature heterodimeric Notch receptor that is then displayed on the cell surface. The binding of Delta, which is displayed on a neighboring cell, triggers the next two proteolytic steps: the complex of Delta and the Notch subunit to which it is bound is endocytosed by the Delta-expressing cell, exposing the extracellular cleavage site in the transmembrane Notch subunit. Note that Notch and Delta interact through their repeated EGF-like domains.

The released Notch tail migrates into the nucleus, where it binds to the Rbpsuh protein and converts it from a transcriptional repressor to a transcriptional activator.

are seven-pass transmembrane proteins that resemble GPCRs in structure. There are seven such receptors in humans. When activated by Wnt binding, Frizzled proteins recruit the scaffold protein **Dishevelled**, which is required for relaying the signal down all three signaling pathways. We focus here on the first pathway.

The **Wnt/β-catenin pathway** acts by regulating the proteolysis of the multifunctional protein **β-catenin** (or *Armadillo* in flies), which functions both in cell–cell adhesion (discussed in Chapter 19) and in gene regulation. In this pathway (but not in the other Wnt pathways), Wnts act by binding to both a Frizzled protein and a co-receptor protein that is related to the low-density lipoprotein (LDL) receptor protein (discussed in Chapter 13) and is therefore called an **LDL-receptor-related protein** (**LRP**). In epithelial cells, most of a cell's β-catenin is located at cell–cell adherens junctions, where it is associated with *cadherins,* which are transmembrane cell–cell adhesion proteins. As discussed in Chapter 19, the β-catenin in these junctions helps link the cadherins to the actin cytoskeleton. In both epithelial and nonepithelial cells, the β-catenin that is not associated with cadherins is rapidly degraded in the cytoplasm.

The degradation of cytoplasmic β-catenin depends on a large protein *degradation complex*, which binds β-catenin and keeps it out of the nucleus while promoting its degradation. The complex contains at least four other proteins: a serine/threonine kinase called *casein kinase 1 (CK1)* phosphorylates the β-catenin on a serine, priming it for further phosphorylation by another serine/threonine kinase called *glycogen synthase kinase 3 (GSK3);* this final phosphorylation marks the protein for ubiquitylation and rapid degradation in proteasomes. Two scaffold proteins called *axin* and *Adenomatous polyposis coli (APC)* hold the protein complex together (**Figure 15–77**A). APC gets its name from the finding that the gene encoding it is often mutated in a type of benign tumor (adenoma) of the colon; the tumor projects into the lumen as a polyp and can eventually become malignant. (This APC is not to be confused with the Anaphase Promoting Complex, or APC, that plays a central part in selective protein degradation during the cell cycle—see Figure 17–20A.)

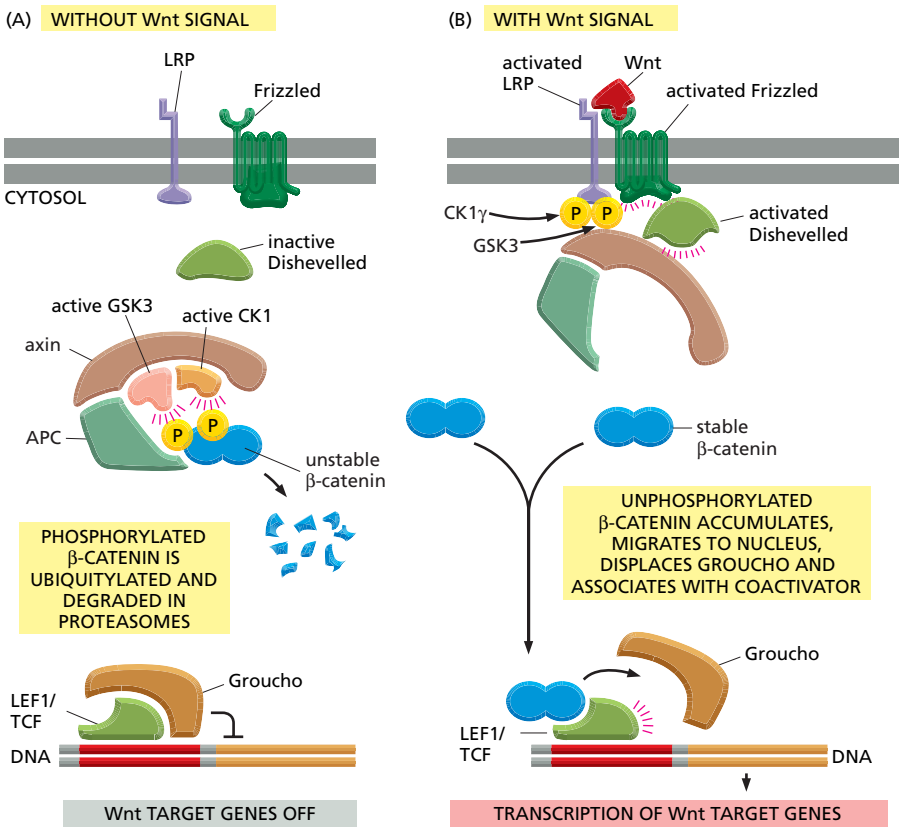


Figure 15–77 The Wnt/β-catenin signaling pathway. (A) In the absence of a Wnt signal, β-catenin that is not bound to the cytosolic tail of cadherin proteins (not shown) becomes bound by a degradation complex containing APC, axin, GSK3, and CK1. In this complex, β-catenin is phosphorylated by CK1 and then by GSK3, triggering its ubiquitylation and degradation in proteasomes. Wnt-responsive genes are kept inactive by the Groucho co-repressor protein bound to the gene regulatory protein LEF1/TCF. (B) Wnt binding to Frizzled and LRP clusters the two types of receptors together, resulting in the recruitment of the degradation complex to the plasma membrane and the phosphorylation of the cytosolic tail of LRP by GSK3 and then by CK1γ. Axin binds to the phosphorylated LRP and is inactivated and/or degraded. The loss of axin from the degradation complex inactivates the complex and thereby blocks β-catenin phosphorylation and ubiquitylation, which allows unphosphorylated β-catenin to accumulate and translocate to the nucleus. Dishevelled and probably a G protein are required for the signaling pathway to operate; both bind to Frizzled and Dishevelled becomes phosphorylated (not shown), but their functional roles are unknown.

Once in In the nucleus, β-catenin binds to LEF1/TCF, displaces the co-repressor Groucho, and acts as a coactivator to stimulate the transcription of Wnt target genes.

The binding of a Wnt protein to both a Frizzled and LRP receptor brings the two receptors together to form a complex. In a poorly understood process, the two protein kinases, GSK3 and then CK1γ, phosphorylate the cytosolic tail of the LRP receptor, allowing the LRP tail to recruit and inactivate axin, thereby disrupting the degradation complex in the cytoplasm. In this way, the phosphorylation and degradation of β-catenin are inhibited, enabling unphosphorylated β-catenin gradually to accumulate and translocate to the nucleus, where it alters the pattern of gene transcription (Figure 15–77B).

In the absence of Wnt signaling, Wnt-responsive genes are kept silent by an inhibitory complex of gene regulatory proteins. The complex includes proteins of the *LEF1/TCF* family bound to a co-repressor protein of the *Groucho* family (see Figure 15–77A). In response to a Wnt signal, β-catenin enters the nucleus and binds to the LEF1/TCF proteins, displacing Groucho. The β-catenin now functions as a coactivator, inducing the transcription of the Wnt target genes (see Figure 15–77B). Thus, as in the case of Notch signaling, Wnt/β-catenin signaling triggers a switch from transcriptional repression to transcriptional activation.

Among the genes activated by β-catenin is *c-Myc,* which encodes a protein (c-Myc) that is a powerful stimulator of cell growth and proliferation (discussed in Chapter 17). Mutations of the *Apc* gene occur in 80% of human colon cancers (discussed in Chapter 20). These mutations inhibit the protein's ability to bind β-catenin, so that β-catenin accumulates in the nucleus and stimulates the transcription of *c-Myc* and other Wnt target genes, even in the absence of Wnt signaling. The resulting uncontrolled cell growth and proliferation promote the development of cancer.

Various secreted inhibitory proteins regulate Wnt signaling in development. Some bind to the LRP receptors and promote their down-regulation, whereas others compete with Frizzled receptors for secreted Wnts. In *Drosophila* at least, Wnts activate negative feedback loops, in which Wnt target genes encode proteins that help shut the response off; some of these proteins inhibit Dishevelled, and others are secreted inhibitors.

Hedgehog Proteins Bind to Patched, Relieving Its Inhibition of Smoothened

Hedgehog proteins and Wnt proteins act in similar ways. Both are secreted signal molecules, which act as local mediators and morphogens in many developing invertebrate and vertebrate tissues. Both proteins are modified by covalently attached lipids, depend on secreted or cell-surface-bound heparan sulfate proteoglycans (discussed in Chapter 19) for their action, and activate latent gene regulatory proteins by inhibiting their proteolysis. They both trigger a switch from transcriptional repression to transcriptional activation, and excessive signaling along either pathway in adult cells can lead to cancer. They even use some of the same intracellular signaling proteins and sometimes collaborate to mediate a response.

The **Hedgehog proteins** were discovered in *Drosophila,* where this protein family has only one member. Mutation of the *Hedgehog* gene produces a larva covered with spiky processes (denticles), like a hedgehog. At least three genes encode Hedgehog proteins in vertebrates—*Sonic, Desert,* and *Indian hedgehog*. The active forms of all Hedgehog proteins are covalently coupled to cholesterol, as well as to a fatty acid chain. The cholesterol is added during an unusual processing step, in which a precursor protein cleaves itself to produce a smaller, cholesterol-containing signal protein. Most of what we know about the downstream signaling pathway activated by Smoothened, however, came initially from genetic studies in flies, and it is the fly pathway that we summarize here.

Three transmembrane proteins—Patched, Smoothened, and iHog—mediate the responses to Hedgehog proteins. **Patched** is predicted to cross the plasma membrane 12 times, and, although much of it is in intracellular vesicles, some is on the cell surface where it binds the Hedgehog protein. **iHog** proteins

have four or five immunoglobulin-like domains and two or three fibronectin-type-III-like domains; they are on the cell surface and are also thought to serve as receptors for Hedgehog proteins, probably acting as co-receptors with Patched. **Smoothened** is a seven-pass transmembrane protein with a structure very similar to a Frizzled protein. In the absence of a Hedgehog signal, Patched through some unknown mechanism, keeps Smoothened sequestered and inactive in intracellular vesicles. The binding of Hedgehog to iHog and Patched inhibits the activity of Patched and induces its endocytosis and degradation. The result is that Smoothened becomes phosphorylated, translocates to the cell surface, and relays the signal downstream.

The downstream effects are mediated by a latent gene regulatory protein called **Cubitus interruptus** (**Ci**). In the absence of a Hedgehog signal, Ci is ubiquitylated and proteolytically cleaved in proteasomes. Instead of being completely degraded, however, Ci is processed to form a smaller protein, which accumulates in the nucleus, where it acts as a transcriptional repressor, helping to keep Hedgehog-responsive genes silent. The proteolytic processing of the Ci protein depends on its phosphorylation by three serine/threonine protein kinases—PKA and two kinases also used in the Wnt pathway, namely GSK3 and CK1. As in the Wnt pathway, the proteolytic processing occurs in a multiprotein complex. The complex includes the serine/threonine kinase *Fused* and a scaffold protein *Costal2*, which stably associates with Ci, recruits the three other kinases, and binds the complex to microtubules, thereby keeping unprocessed Ci out of the nucleus (**Figure 15–78**A).

When the Hedgehog pathway is activated and Smoothened is thereby let loose in the plasma membrane, it recruits the protein complexes containing Ci, Fused, and Costal2. Costal2 is no longer able to bind the other three kinases, and so Ci is no longer cleaved. Unprocessed Ci protein can now enter the nucleus and activate the transcription of Hedgehog target genes (Figure 15–78B). Among the genes activated by Ci is *Patched* itself; the resulting increase in Patched protein on the cell surface inhibits further Hedgehog signaling—providing yet another example of a negative feedback loop.

Many gaps in the understanding of the Hedgehog signaling pathway remain to be filled in. It is not known, for example, how Patched keeps Smoothened inactive and intracellular. As the structure of Patched resembles a transmembrane transporter protein, it has been proposed that it may transport a small molecule into the cell that keeps Smoothened sequestered in vesicles.

Even less is known about the more complex Hedgehog pathway in vertebrate cells. In addition to there being at least three types of vertebrate Hedgehog proteins, there are three Ci-like gene regulatory proteins (*Gli1, Gli2,* and *Gli3*) downstream of Smoothened. Only Gli3 has been shown to undergo proteolytic processing like Ci and to act as either a transcriptional repressor or a transcriptional activator. Both Gli1 and Gli2 are thought to act exclusively as transcriptional activators. Moreover, in vertebrates, Smoothened, upon activation, becomes localized to a very specific site in the plasma membrane—the surface of the primary cilium, which projects from the surface of most types of vertebrate cells (as discussed in Chapter 16). The primary cilium thus acts as a Hedgehog signaling center, and the Gli proteins are also concentrated here. This arrangement presumably increases the speed and efficiency of the signaling process.

Hedgehog signaling can promote cell proliferation, and excessive Hedgehog signaling can lead to cancer. Inactivating mutations in one of the two human *Patched* genes, for example, which leads to excessive Hedgehog signaling, occurs frequently in *basal cell carcinoma* of the skin, the most common form of cancer in Caucasians. A small molecule called *cyclopamine,* made by a meadow lily, is being used to treat cancers associated with excessive Hedgehog signaling. It blocks Hedgehog signaling by binding tightly to Smoothened and inhibiting its activity. It was originally identified because it causes severe developmental defects in the progeny of sheep grazing on such lilies; these include the presence of a single central eye (a condition called *cyclopia),* which is also seen in mice that are deficient in Hedgehog signaling.

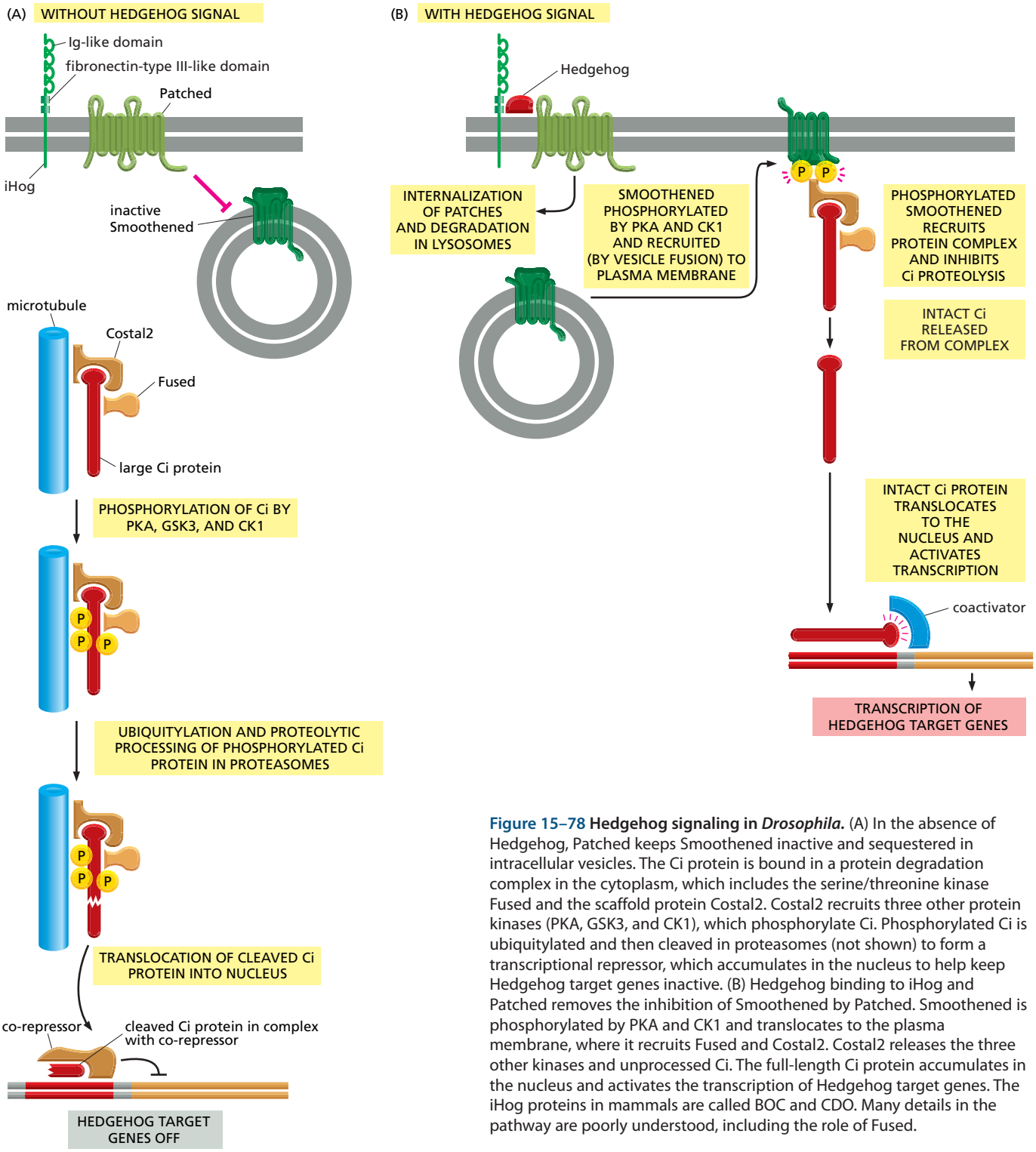


Figure 15–78 Hedgehog signaling in *Drosophila*. (A) In the absence of Hedgehog, Patched keeps Smoothened inactive and sequestered in intracellular vesicles. The Ci protein is bound in a protein degradation complex in the cytoplasm, which includes the serine/threonine kinase Fused and the scaffold protein Costal2. Costal2 recruits three other protein kinases (PKA, GSK3, and CK1), which phosphorylate Ci. Phosphorylated Ci is ubiquitylated and then cleaved in proteasomes (not shown) to form a transcriptional repressor, which accumulates in the nucleus to help keep Hedgehog target genes inactive. (B) Hedgehog binding to iHog and Patched removes the inhibition of Smoothened by Patched. Smoothened is phosphorylated by PKA and CK1 and translocates to the plasma membrane, where it recruits Fused and Costal2. Costal2 releases the three other kinases and unprocessed Ci. The full-length Ci protein accumulates in the nucleus and activates the transcription of Hedgehog target genes. The iHog proteins in mammals are called BOC and CDO. Many details in the pathway are poorly understood, including the role of Fused.

Many Stressful and Inflammatory Stimuli Act Through an NFκB-Dependent Signaling Pathway

The **NFκB proteins** are latent gene regulatory proteins that are present in most animal cells and are central to many stressful, inflammatory, and innate immune responses. These responses occur as a reaction to infection or injury and help protect stressed multicellular organisms and their cells (discussed in Chapter 24). An excessive or inappropriate inflammatory response in animals can also damage tissue and cause severe pain, and chronic inflammation can lead to cancer; as in the case of Wnt and Hedgehog signaling, excessive NFκB

signaling is found in a number of human cancers. NFκB proteins also have important roles during normal animal development. The *Drosophila* NFκB family member *Dorsal*, for example, has a crucial role in specifying the dorsal–ventral axis of the developing fly embryo (discussed in Chapter 22).

Various cell-surface receptors activate the NFκB signaling pathway in animal cells. *Toll receptors* in *Drosophila* and *Toll-like receptors* in vertebrates, for example, recognize pathogens and activate this pathway in triggering innate immune responses (discussed in Chapter 24). The receptors for *tumor necrosis factor α (TNFα)* and *interleukin-1 (IL1)*, which are vertebrate cytokines especially important in inducing inflammatory responses, also activate this pathway. The Toll, Toll-like, and IL1 receptors belong to the same family of proteins, whereas TNF receptors belong to a different family; all of them, however, act in similar ways to activate NFκB. When activated, they trigger a multiprotein ubiquitylation and phosphorylation cascade that releases NFκB from an inhibitory protein complex, so that it can translocate to the nucleus and turn on the transcription of hundreds of genes that participate in inflammatory and innate immune responses.

There are five NFκB proteins in mammals (*RelA*, *RelB*, *c-Rel*, *NFκB1*, and *NFκB2*), and they form a variety of homodimers and heterodimers, each of which activates its own characteristic set of genes. Inhibitory proteins called **IκB** bind tightly to the dimers and hold them in an inactive state within the cytoplasm of unstimulated cells. There are three major IκB proteins in mammals (IκB α, β, and ε), and the signals that release NFκB dimers do so by triggering a signaling pathway that leads to the phosphorylation, ubiquitylation, and consequent degradation of the IκB proteins. The phosphorylation of IκB is mediated by *IκB kinase (IKK)*, which is a multiprotein complex containing two serine/threonine protein kinases (IKKα and IKKβ) and a regulatory protein called *NEMO* (for NFκB essential modifier), or IKKγ (**Figure 15–79**).

Among the genes activated by the released NFκB is the gene that encodes IκBα, one of three IκB isoforms that holds NFκB inactive in the cytosol of resting cells. This activation leads to the resynthesis of IκBα protein, which binds to NFκB and inactivates it, creating a negative feedback loop (**Figure 15–80**A). Experiments on TNFα-induced responses, as well as computer modeling studies of the responses, indicate that the negative feedback produces two types of

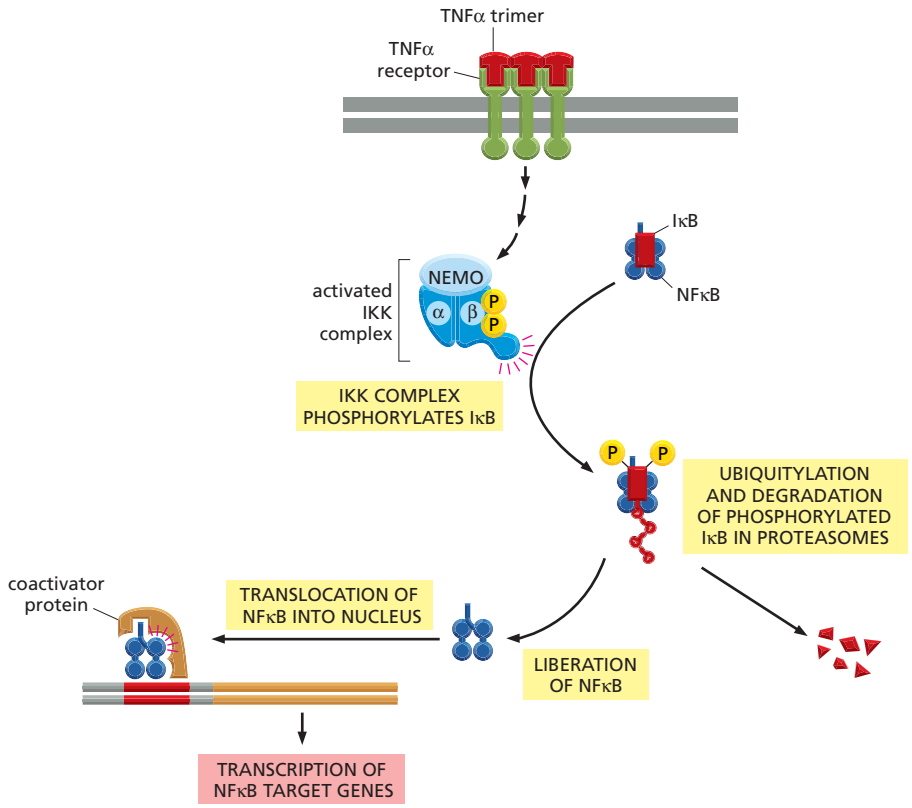
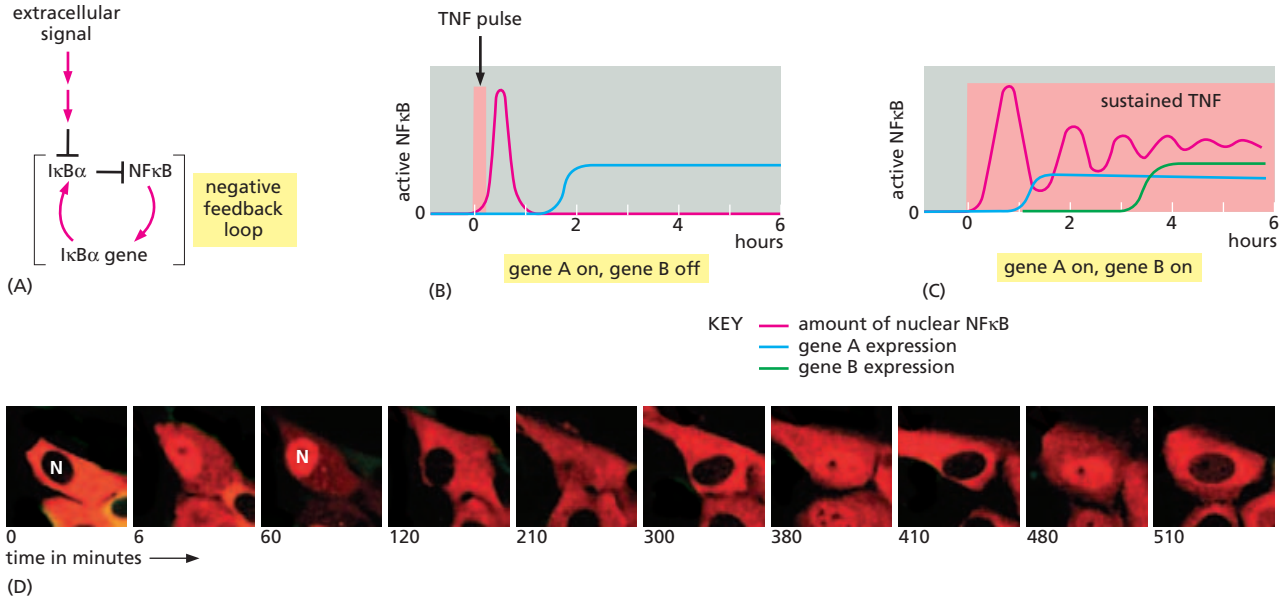


Figure 15–79 The activation of the NFκB pathway by TNFα. Both TNFα and its receptors are trimers. The binding of TNFα causes a rearrangement of the clustered cytosolic tails of the receptors, which now recruit various signaling proteins, resulting in the activation of a serine/threonine protein kinase that phosphorylates and activates IκB kinase kinase (IKK). IKK is a heterotrimer composed of two kinase subunits (IKKα and IKKβ) and a regulatory subunit called NEMO. IKKβ then phosphorylates IκB on two serines, which marks the protein for ubiquitylation and degradation in proteasomes. The released NFκB translocates into the nucleus, where, in collaboration with coactivator proteins, it stimulates the transcription of its target genes.



NFκB responses, depending on the duration of the TNFα stimulus. Short exposure (less than an hour) to TNFα produces a short period of NFκB activation that is independent of the duration of the TNFα stimulus; the negative feedback through IκBα shuts off the response after about an hour. Prolonged exposure, by contrast, produces slow oscillations of NFκB activation, in which activation is followed by IκBα-mediated inactivation, which is followed by IκBα destruction and NFκB reactivation, and so on; the oscillations can persist for several hours before fading, even when the stimulus is maintained. Importantly, the two types of responses induce different patterns of gene expression, as some NFκB target genes turn on only in response to the prolonged oscillatory activation of NFκB (Figure 15–80B, C, and D). The negative feedback through IκBα is required for both types of responses: in cells deficient in IκBα, even a short exposure to TNFα induces a sustained activation of NFκB, without oscillations, and all of the NFκB-responsive genes are activated.

Thus far, we have discussed cell signaling mainly in animals, with a few diversions into yeasts and bacteria. But intercellular signaling is just as important for plants as it is for animals, although the mechanisms and molecules used are mainly different, as we discuss next.

Figure 15–80 Negative feedback in the NFκB signaling pathway induces oscillations in NFκB activation. (A) Drawing showing how activated NFκB stimulates the transcription of IκBα, the protein product of which acts back in the cytoplasm to sequester NFκB there; if the stimulus is persistent, the newly made IκBα protein will then be ubiquitylated and degraded, liberating active NFκB again so that it can return to the nucleus and activate transcription (see Figure 15–79). (B) A short exposure to TNFα produces a single, short pulse of NFκB activation, beginning within minutes and ending by 1 hour. This response turns on the transcription of gene A but not gene B. (C) A sustained exposure to TNFα for the entire 6 hours of the experiment produces oscillations in NFκB activation that damp down over time. This response turns on the transcription of both genes; gene B turns on only after several hours, indicating that gene B transcription requires prolonged activation of NFκB, for reasons that are not understood. (D) These time-lapse confocal fluorescence micrographs from a different study of TNFα stimulation show the oscillations of NFκB in a cultured cell, as indicated by the periodic movement into the nucleus (N) of a fusion protein composed of NFκB fused to a red fluorescent protein. In the cell at the upper part of the micrographs, NFκB is active and in the nucleus at 6, 60, 210, 380, and 480 minutes, but it is exclusively in the cytoplasm at 0, 120, 300, 410, and 510 minutes. (A–C, based on data from A. Hoffmann et al., *Science* 298:1241–1245, 2002, and adapted from A.Y. Ting and D. Endy, *Science* 298:1189–1190, 2002; D, from D.E. Nelson et al., *Science* 306:704–708, 2004. All with permission from AAAS.)

Summary

Some signaling pathways that are especially important in animal development depend on proteolysis to control the activity and location of latent gene regulatory proteins. Notch receptors are themselves such latent gene regulatory proteins, which are activated by cleavage when Delta (or a related ligand) on another cell binds to them; the cleaved cytosolic tail of Notch migrates into the nucleus, where it stimulates the transcription of Notch-responsive genes. In the Wnt/β-catenin signaling pathway, by contrast, the proteolysis of the latent gene regulatory protein β-catenin is inhibited when a secreted Wnt protein binds to both a Frizzled and LRP receptor protein; as a result, β-catenin accumulates in the nucleus and activates the transcription of Wnt target genes.

Hedgehog signaling in flies works much like Wnt signaling. In the absence of a signal, a bifunctional, cytoplasmic gene regulatory protein, Ci, is proteolytically cleaved to form a transcriptional repressor that keeps Hedgehog target genes silenced. The binding of Hedgehog to its receptors (Patched and iHog) inhibits the proteolytic processing of Ci; as a result, the larger form of Ci accumulates in the nucleus and activates the transcription of Hedgehog-responsive genes. In Notch, Wnt, and Hedgehog signaling, the extracellular signal triggers a switch from transcriptional repression to transcriptional activation.

Signaling through the latent gene regulatory protein NFκB also depends on proteolysis. NFκB proteins are normally held in an inactive state by inhibitory IκB proteins

in the cytoplasm. A variety of extracellular stimuli, including proinflammatory cytokines, trigger the phosphorylation and ubiquitylation of IκB, marking it for degradation; this enables the NFκB to translocate to the nucleus and activate the transcription of its target genes. NFκB also activates the transcription of the gene that encodes IκBα, creating a negative feedback loop, which can produce prolonged oscillations in NFκB activity with sustained extracellular signaling.

SIGNALING IN PLANTS

In plants, as in animals, cells are in constant communication with one another. Plant cells communicate to coordinate their activities in response to the changing conditions of light, dark, and temperature, which guide the plant's cycle of growth, flowering, and fruiting. Plant cells also communicate to coordinate activities in their roots, stems, and leaves. In this final section, we consider how plant cells signal to one another and how they respond to light. Less is known about the receptors and intracellular signaling mechanisms involved in cell communication in plants than is known in animals, and we will concentrate mainly on how the receptors and mechanisms differ from those used by animals. We discuss some of the details of plant development in Chapter 22.

Multicellularity and Cell Communication Evolved Independently in Plants and Animals

Although plants and animals are both eucaryotes, they have evolved separately for more than a billion years. Their last common ancestor is thought to have been a unicellular eucaryote that had mitochondria but no chloroplasts; the plant lineage acquired chloroplasts after plants and animals diverged. The earliest fossils of multicellular animals and plants date from almost 600 million years ago. Thus, it seems that plants and animals evolved multicellularity independently, each starting from a different unicellular eucaryote, sometime between 1.6 and 0.6 billion years ago (**Figure 15–81**).

If multicellularity evolved independently in plants and animals, the molecules and mechanisms used for cell communication will have evolved separately and would be expected to be different. There should be some degree of resemblance, however, because the genes in both plants and animals diverged from those contained by their last common unicellular ancestor. Thus, whereas both plants and animals use nitric oxide, cyclic GMP, Ca^{2+}, and Rho family GTPases for signaling, there are no homologs of the nuclear receptor family, Ras, JAK, STAT, TGFβ, Notch, Wnt, or Hedgehog encoded by the completely sequenced genome of *Arabidopsis thaliana,* the small flowering plant. Similarly, plants do not seem to use cyclic AMP for intracellular signaling.

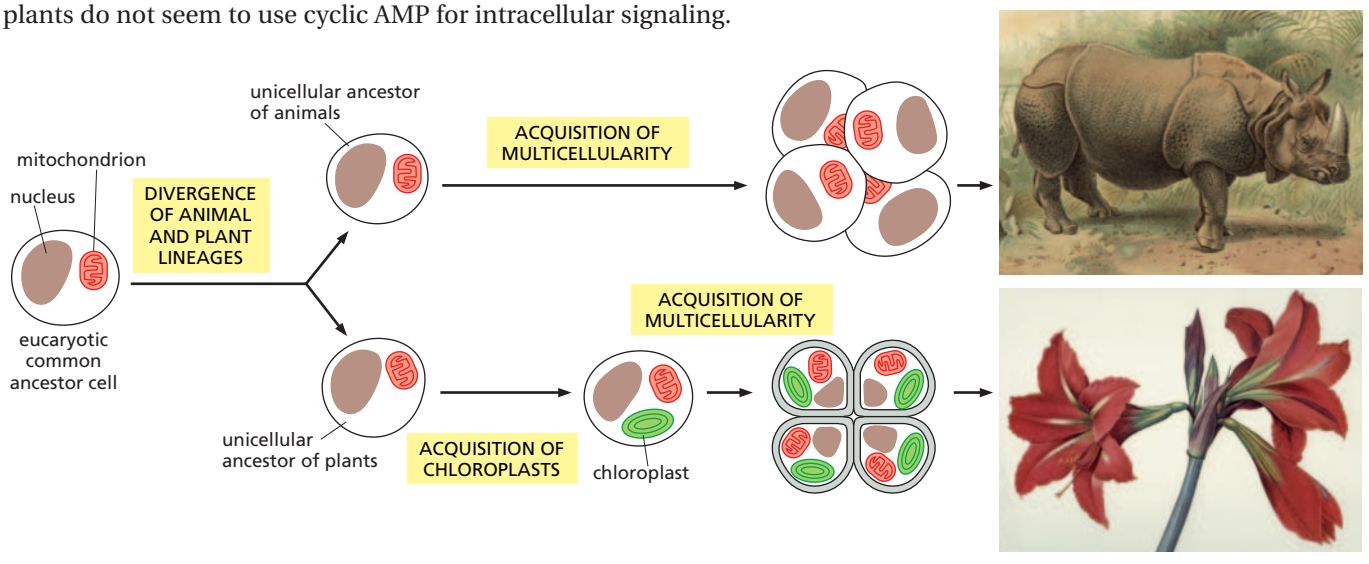


Figure 15–81 The proposed divergence of plant and animal lineages from a common unicellular eucaryotic ancestor. The plant lineage acquired chloroplasts after the two lineages diverged. Both lineages independently gave rise to multicellular organisms—plants and animals. (Paintings courtesy of John Innes Foundation.)

Much of what is known about the molecular mechanisms involved in signaling in plants has come from genetic studies on *Arabidopsis*. Although the specific molecules used in cell communication in plants often differ from those used in animals, the general strategies are frequently very similar. Both, for example, use enzyme-coupled cell-surface receptors, as we now discuss.

Receptor Serine/Threonine Kinases Are the Largest Class of Cell-Surface Receptors in Plants

Whereas most cell-surface receptors in animals are G-protein-coupled (GPCRs), most found so far in plants are enzyme-coupled. Moreover, whereas the largest class of enzyme-coupled receptors in animals is the receptor tyrosine kinase (RTK) class, this type of receptor is extremely rare in plants. Plants do, however, have many cytoplasmic tyrosine kinases, and tyrosine phosphorylation and dephosphorylation have important roles in plant cell signaling. Instead of RTKs, plants rely largely on a great diversity of transmembrane *receptor serine/threonine kinases*. Although very different from the corresponding animal receptors in most respects, they resemble them in having a typical serine/threonine kinase cytoplasmic domain and an extracellular ligand-binding domain. The most abundant types of these receptors have a tandem array of extracellular leucine-rich repeats (**Figure 15–82**) and are therefore called **leucine-rich repeat (LRR) receptor kinases**.

There are about 175 LRR receptor kinases encoded by the *Arabidopsis* genome. One of the best characterized is the *Clavata1/Clavata2 (Clv1/Clv2) receptor* complex. Mutations that inactivate either of the two receptor subunits cause the production of flowers with extra floral organs and a progressive enlargement of both the shoot and floral *meristems,* the groups of self-renewing stem cells that produce the cells that give rise to stems, leaves, and flowers (discussed in Chapter 22). The extracellular signal molecule that binds to the receptor is thought to be a small protein called *Clv3,* which is secreted by neighboring cells. The binding of Clv3 to the Clv1/Clv2 receptor suppresses meristem growth, either by inhibiting cell division there or, more probably, by stimulating cell differentiation (**Figure 15–83**A).

The intracellular signaling pathway from the Clv1/Clv2 receptor to the cell response is largely unknown, but it includes a serine/threonine protein phosphatase that inhibits the signaling. Other signaling proteins in the pathway include a Rho family GTPase and a nuclear gene regulatory protein that is distantly related to animal homeodomain proteins. Mutations that inactivate this gene regulatory protein have the opposite effect from those that inactivate the Clv1/Clv2 receptor: cell division is greatly decreased in the shoot meristem, and the plant produces flowers with too few organs. Thus, the intracellular signaling pathway activated by the Clv1/Clv2 receptor seems to stimulate cell differentiation by inhibiting the gene regulatory protein that normally inhibits cell differentiation (Figure 15–83B).

A different LRR receptor kinase in *Arabidopsis*, called *Bri1*, is part of a cell-surface steroid hormone receptor. Plants synthesize a class of steroids that are

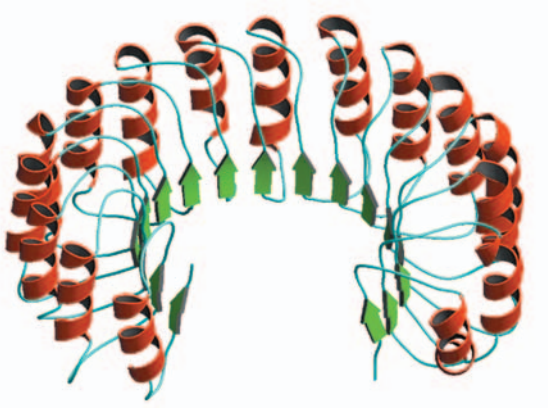


Figure 15–82 The three-dimensional structure of leucine-rich repeats, similar to those found in the LRR serine/threonine receptor kinases. Multiple copies of such repeats are present in the extracellular domain of LRR receptor kinases, where they participate in binding the signal molecule. (Courtesy of David Lawson.)

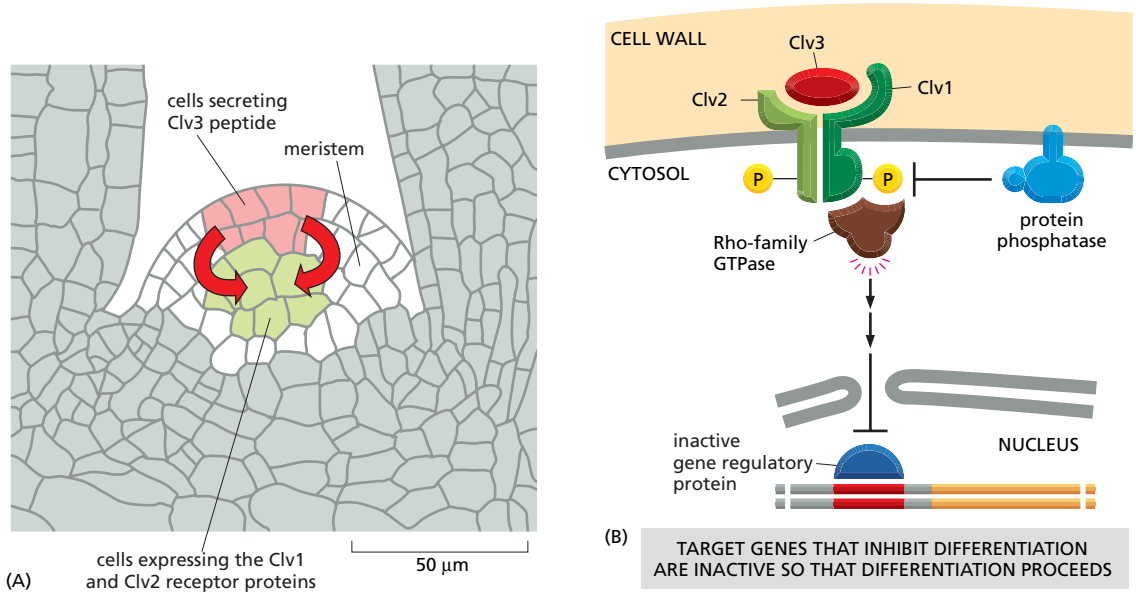


Figure 15–83 A hypothetical model for how Clv3 and the Clv1/Clv2 receptor regulate cell proliferation and/or differentiation in the shoot meristem. (A) Cells in the outer layer of the meristem *(light red)* secrete Clv3 peptide, which binds to the Clv1/Clv2 receptor protein on target cells in an adjacent, more central region of the meristem *(green),* stimulating the differentiation of the target cells. (B) Some parts of the intracellular signaling pathway activated by Clv3 binding. When activated by Clv3 binding, Clv1 phosphorylates the receptor proteins on serines and threonines, thereby activating the receptor complex and leading to the activation of a Rho family GTPase. The signaling pathway after this point is unclear, but it leads to the inhibition of a gene regulatory protein in the nucleus called Wuschel. Because Wuschel normally blocks the transcription of genes required for differentiation, its inhibition of Clv3 signaling allows differentiation to proceed. The protein phosphatase dephosphorylates the receptor proteins and thereby negatively regulates the signaling pathway.

called **brassinosteroids** because they were originally identified in the mustard family Brassicaceae, which includes *Arabidopsis*. These plant signal molecules regulate the growth and differentiation of plants throughout their life cycle. Binding of a brassinosteroid to a Bri1 cell-surface receptor kinase initiates a signaling cascade that uses a GSK3 protein kinase and a protein phosphatase to regulate the phosphorylation and degradation of specific gene regulatory proteins in the nucleus, and thereby specific gene transcription. Mutant plants that are deficient in the Bri1 receptor kinase are insensitive to brassinosteroids and are therefore dwarfs.

The LRR receptor kinases are only one of many classes of transmembrane receptor serine/threonine kinases in plants. There are at least six additional families, each with its own characteristic set of extracellular domains. The *lectin receptor kinases,* for example, have extracellular domains that bind carbohydrate signal molecules. The *Arabidopsis* genome encodes over 300 receptor serine/threonine kinases, which makes them the largest family of receptors known in plants. Many are involved in defense responses against pathogens.

Ethylene Blocks the Degradation of Specific Gene Regulatory Proteins in the Nucleus

Various **growth regulators** (also called **plant hormones**) help to coordinate plant development. They include *ethylene, auxin, cytokinins, gibberellins,* and *abscisic acid,* as well as brassinosteroids. Growth regulators are all small molecules made by most plant cells. They diffuse readily through cell walls and can either act locally or be transported to influence cells further away. Each growth regulator can have multiple effects. The specific effect depends on environmental conditions, the nutritional state of the plant, the responsiveness of the target cells, and which other growth regulators are acting.

Ethylene is an important example. This small gas molecule (**Figure 15–84**A) can influence plant development in various ways; it can, for example, promote fruit ripening, leaf abscission, and plant senescence. It also functions as a stress signal in response to wounding, infection, flooding, and so on. When the shoot of a germinating seedling, for instance, encounters an obstacle, such as a piece of gravel underground in the soil, the seedling responds to the encounter in three ways. First, it thickens its stem, which can then exert more force on the obstacle. Second, it shields the tip of the shoot by increasing the curvature of a specialized hook structure. Third, it reduces the shoot's tendency to grow away from the direction of gravity, so as to avoid the obstacle. This *triple response* is controlled by ethylene (Figure 15–84B and C).

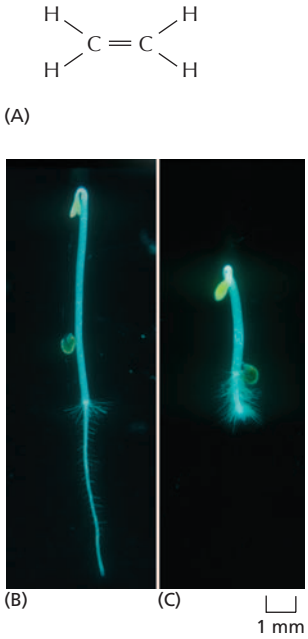


Figure 15–84 The ethylene-mediated triple response that occurs when the growing shoot of a germinating seedling encounters an obstacle underground. (A) The structure of ethylene. (B) Before the encounter, the shoot is growing upward and is long and thin. (C) After the encounter, the shoot thickens, and the protective hook (at *top*) increases its curvature to protect the tip of the shoot. The shoot also alters its direction of growth to grow around the obstacle (not shown). (Courtesy of Melanie Webb.)

Plants have various ethylene receptors, which are located in the endoplasmic reticulum and are all structurally related. They are dimeric, multipass transmembrane proteins, with a copper-containing ethylene-binding domain and a domain that interacts with a protein called *CTR1*, which is closely related in sequence to the Raf MAP kinase kinase kinase discussed earlier (see Figure 15–60). The function of CTR1 in ethylene signaling depends on its serine/threonine kinase activity and the association of its N-terminal domain with the ethylene receptors. Surprisingly, it is the empty receptors that are active and keep CTR1 active. By an unknown signaling mechanism, active CTR1 stimulates the ubiquitylation and degradation in proteasomes of a nuclear gene regulatory protein called *EIN3*, which is required for the transcription of ethylene-responsive genes. In this way, the empty but active receptors and active CTR1 keep ethylene-response genes off. The EIN protein gets its name from the finding that plants with inactivating mutations in the gene that encodes it are ethylene insensitive.

Ethylene binding inactivates the receptors, altering their conformation so that they no longer bind to CTR1. As a result, CTR1 is inactivated, and the downstream signaling pathway emanating from it is blocked; the EIN3 protein is no longer ubiquitylated and degraded and can now activate the transcription of the large number of ethylene-responsive genes (**Figure 15–85**).

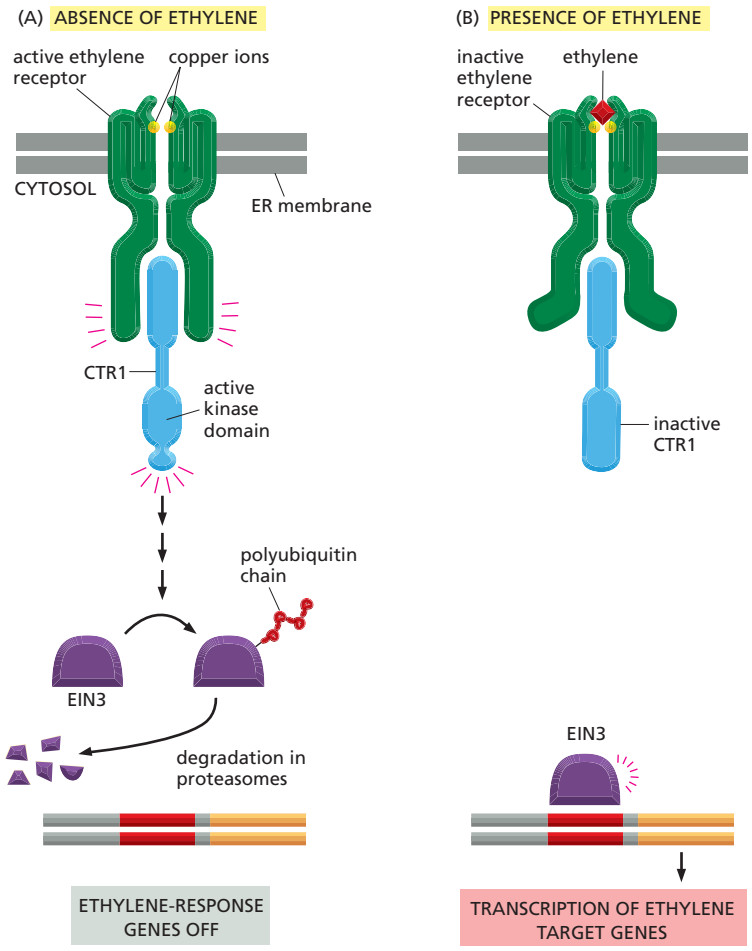


Figure 15–85 A current view of the ethylene signaling pathway. (A) In the absence of ethylene, both the receptors and CTR1 are active, causing the ubiquitylation and destruction of the EIN3 protein, the gene regulatory protein in the nucleus that is responsible for the transcription of ethylene-responsive genes. (B) The binding of ethylene inactivates the receptors and disrupts the interaction between the receptors and CTR1. The EIN3 protein is not degraded and can therefore activate the transcription of ethylene-responsive genes.

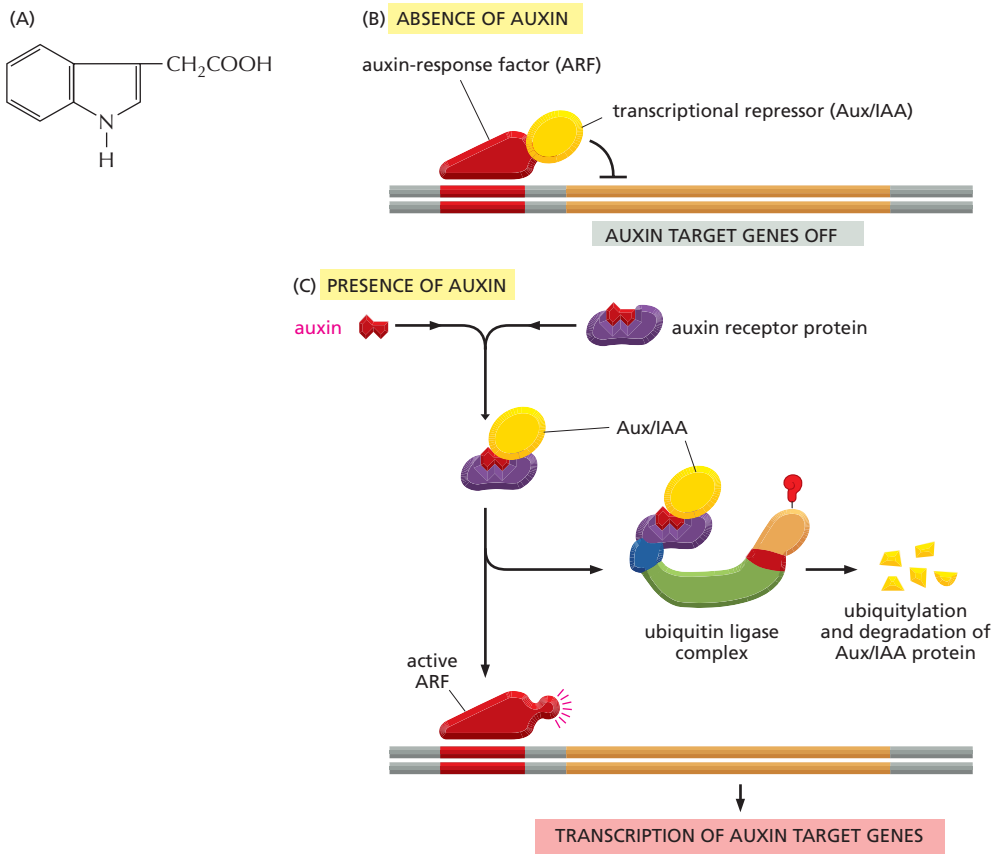


Figure 15–86 The auxin signaling pathway. (A) The structure of the auxin indole-3-acetic acid. (B) In the absence of auxin, a transcriptional repressor protein (called Aux/IAA) binds and suppresses a gene regulatory protein (called auxin-response factor, ARF), which is required for the transcription of auxin-responsive genes. (C) The auxin receptor proteins are mainly located in the nucleus and form part of ubiquitin ligase complexes (not shown). When activated by auxin binding, the receptor–auxin complexes recruit the ubiquitin ligase complexes, which ubiquitylate the Aux/IAA proteins, marking them for degradation in proteasomes. ARF is now free to activate the transcription of auxin-responsive genes. There are many ARFs, Aux/IAA proteins, and auxin receptors that work as illustrated.

A somewhat different strategy is involved in the regulation of *auxin*-responsive genes. Moreover, the way that auxin controls the direction and pattern of plant growth is unlike any mechanism observed in animals, as we now discuss.

Regulated Positioning of Auxin Transporters Patterns Plant Growth

The plant hormone **auxin**, which is generally indole-3-acetic acid (**Figure 15–86**A), binds to receptor proteins in the nucleus. It helps plants grow toward light, grow upward rather than branch out, and grow their roots down. It also regulates organ initiation and positioning and helps plants flower and bear fruit. Like ethylene, it influences gene expression by controlling the degradation of gene regulatory proteins in the nucleus, but instead of blocking the ubiquitylation and degradation of gene regulatory proteins required for the expression of auxin-responsive genes, it stimulates the ubiquitylation and degradation of repressor proteins that block the transcription of these genes in unstimulated cells (Figure 15–86B and C).

Auxin is unique in the way that it is transported. Unlike animal hormones, which are usually secreted by a specific endocrine organ and transported to target cells via the circulatory system, auxin has its own transport system. Specific plasma-membrane-bound *influx transporter proteins* and *efflux transporter proteins* move auxin into and out of plant cells, respectively. Distinct gene families encode the influx transporters and efflux transporters, and the two families of proteins are regulated independently. The efflux transporters are composed of *Pin proteins,* and cells can distribute them asymmetrically in the plasma membrane to make the efflux of auxin directional. A row of cells with their auxin efflux transporters confined to the basal plasma membrane, for example, will transport auxin from the top of the plant to the bottom.

In some regions of the plant, the localization of the auxin transporters, and therefore the direction of auxin flow, is highly dynamic and regulated. A cell can

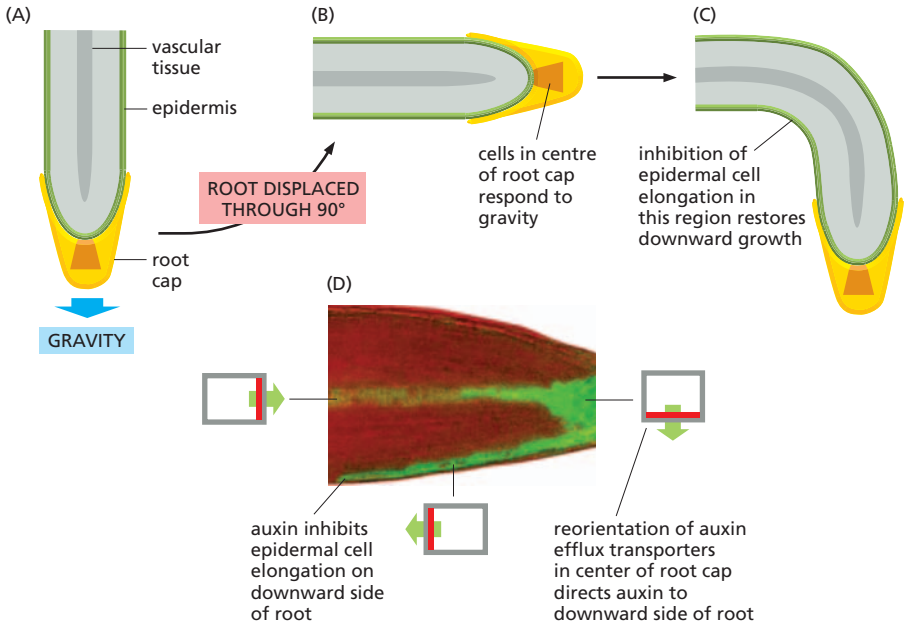


Figure 15–87 Auxin transport and root gravitropism. (A–C) Roots respond to a 90° change in the gravity vector and adjust their direction of growth so that they grow downward again. The cells that respond to gravity are in the center of the root cap, while it is the epidermal cells further back (on the lower side) that decrease their rate of elongation to restore downward growth. (D) The gravity-responsive cells in the root cap redistribute their auxin efflux transporters in response to the displacement of the root. This redirects the auxin flux mainly to the lower part of the displaced root, where it inhibits the elongation of the epidermal cells. The resulting asymmetrical distribution of auxin in the *Arabidopsis* root tip shown here is assessed indirectly, using an auxin-responsive reporter gene that encodes a protein fused to green fluorescent protein (GFP); the epidermal cells on the downward side of the root are green, whereas those on the upper side are not, reflecting the asymmetrical distribution of auxin. The distribution of auxin efflux transporters in the plasma membrane of cells in different regions of the root (shown as *gray rectangles)* is indicated in *red,* and the direction of auxin efflux is indicated by a *green arrow.* (The fluorescence photograph in D is from T. Paciorek et al., *Nature* 435:1251–1256, 2005. With permission from Macmillan Publishers Ltd.)

rapidly redistribute transporters by controlling the traffic of vesicles containing them. The auxin efflux transporters, for example, normally recycle between intracellular vesicles and the plasma membrane. A cell can redistribute these transporters on its surface by inhibiting their endocytosis in one domain of the plasma membrane, causing the transporters to accumulate there. One example occurs in the root, where gravity influences the direction of growth. The auxin efflux transporters are normally distributed symmetrically in the cap cells of the root. Within minutes of a change in the direction of the gravity vector, however, the efflux transporters redistribute to one side of the cells, so that auxin is pumped out toward the side of the root pointing downward. Because auxin inhibits root cell elongation, this redirection of auxin transport causes the root tip to reorient, so that it grows downward again (**Figure 15–87**).

In shoot apical meristems, the distribution of an auxin efflux transporter is also dynamic and regulated. Here, the directional transport of auxin helps determine the regular arrangement of leaves and flowers (see Figure 22–122).

Phytochromes Detect Red Light, and Cryptochromes Detect Blue Light

Plant development is greatly influenced by environmental conditions. Unlike animals, plants cannot move on when conditions become unfavorable; they have to adapt or they die. The most important environmental influence on plants is light, which is their energy source and has a major role throughout their entire life cycle—from germination, through seedling development, to flowering and senescence. Plants have thus evolved a large set of light-sensitive proteins to monitor the quantity, quality, direction, and duration of light. These are usually referred to as *photoreceptors*. However, because the term photoreceptor is also used for light-sensitive cells in the animal retina (see Figure 15–48), we shall use the term *photoprotein* instead.

All photoproteins sense light by means of a covalently attached light-absorbing chromophore, which changes its shape in response to light and then induces a change in the protein's conformation.

The best-known plant photoproteins are the **phytochromes**, which are present in all plants and in some algae but are absent in animals. These are dimeric, cytoplasmic serine/threonine kinases, which respond differentially and reversibly to red and far-red light: whereas red light usually activates the kinase

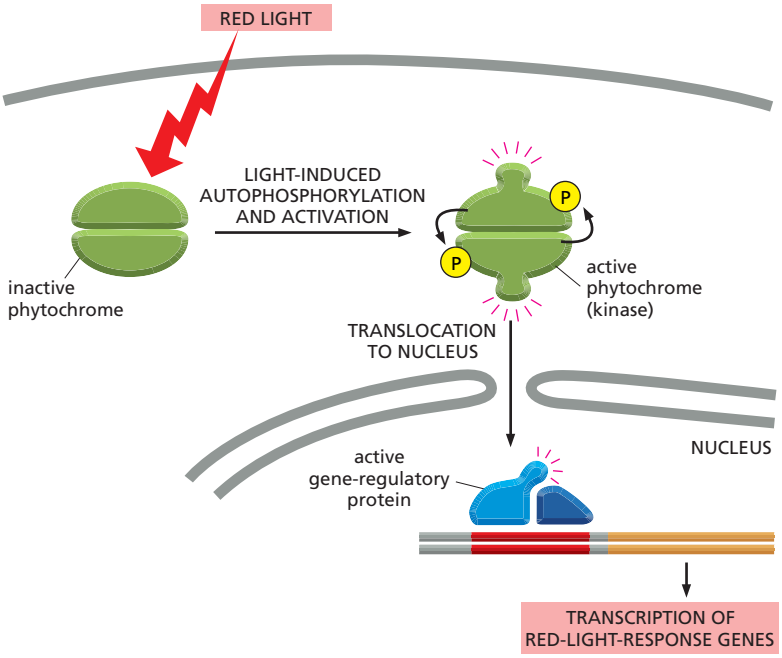


Figure 15–88 A current view of one way in which phytochromes mediate a light response in plant cells. When activated by red light, the phytochrome, which is a dimeric protein kinase, phosphorylates itself and then moves into the nucleus, where it activates gene regulatory proteins to stimulate the transcription of red-light-responsive genes.

activity of the phytochrome, far-red light inactivates it. When activated by red light, the phytochrome is thought to phosphorylate itself and then to phosphorylate one or more other proteins in the cell. In some light responses, the activated phytochrome translocates into the nucleus, where it activates gene regulatory proteins to alter gene transcription (**Figure 15–88**). In other cases, the activated phytochrome activates a latent gene regulatory protein in the cytoplasm, which then translocates into the nucleus to regulate gene transcription. In still other cases, the photoprotein triggers signaling pathways in the cytosol that alter the cell's behavior without involving the nucleus.

Although the phytochromes possess serine/threonine kinase activity, parts of their structure resemble the histidine kinases involved in bacterial chemotaxis discussed earlier. This finding suggests that the plant phytochromes originally evolved from bacterial histidine kinases and only later in evolution altered their substrate specificity from histidine to serine and threonine.

Plants sense blue light using photoproteins of two other sorts, phototropin and cryptochromes. **Phototropin** is associated with the plasma membrane and is partly responsible for *phototropism,* the tendency of plants to grow toward light. Phototropism occurs by directional cell elongation, which is stimulated by auxin, but the links between phototropin and auxin are unknown.

Cryptochromes are flavoproteins that are sensitive to blue light. They are structurally related to blue-light-sensitive enzymes called *photolyases,* which are involved in the repair of ultraviolet-induced DNA damage in all organisms, except most mammals. Unlike phytochromes, cryptochromes are also found in animals, where they have an important role in circadian clocks, which operate in most cells and cycle with a 24-hour rhythm (discussed in Chapter 7). Although cryptochromes are thought to have evolved from the photolyases, they do not have a role in DNA repair.

In this chapter, we have discussed how extracellular signals influence cell behavior. One crucial intracellular target of these signals is the cytoskeleton, which determines cell shape and is responsible for cell movements, as we discuss in the next chapter.

Summary

Plants and animals are thought to have evolved multicellularity and cell communication mechanisms independently, each starting from a different unicellular eucaryote,

which in turn evolved from a common unicellular eucaryotic ancestor. Not surprisingly, therefore, the mechanisms used to signal between cells in animals and in plants have both similarities and differences. Whereas animals rely mainly on GPCRs, for example, plants rely mainly on enzyme-coupled receptors of the receptor serine/threonine kinase type, especially ones with extracellular leucine-rich repeats. Various plant hormones, or growth regulators, including ethylene and auxin, help coordinate plant development. Ethylene acts through intracellular receptors to stop the degradation of specific nuclear gene regulatory proteins, which can then activate the transcription of ethylene-responsive genes. The receptors for some other plant hormones, including auxin, also regulate the degradation of specific gene regulatory proteins, although the details vary. Auxin signaling is unusual in that it has its own highly regulated transport system, in which the dynamic positioning of plasma-membrane-bound auxin transporters controls the direction of auxin flow and thereby the direction of plant growth. Light has an important role in regulating plant development. These light responses are mediated by a variety of light-sensitive photoproteins, including phytochromes, which are responsive to red light, and cryptochromes and phototropin, which are sensitive to blue light.

PROBLEMS

Which statements are true? Explain why or why not.

15–1 The receptors involved in paracrine, synaptic, and endocrine signaling all have very high affinity for their respective signaling molecules.

15–2 All small intracellular mediators (second messengers) are water soluble and diffuse freely through the cytosol.

15–3 In the regulation of molecular switches, protein kinases and guanine nucleotide exchange factors (GEFs) always turn proteins on, whereas protein phosphatases and GTPase-activating proteins (GAPs) always turn proteins off.

15–4 In contrast to the more direct signaling pathways used by nuclear receptors, catalytic cascades of intracellular mediators provide numerous opportunities for amplifying the responses to extracellular signals.

15–5 Binding of extracellular ligands to receptor tyrosine kinases activates the intracellular catalytic domain by propagating a conformational change across the lipid bilayer through a single transmembrane α helix.

15–6 Protein tyrosine phosphatases display exquisite specificity for their substrates, unlike serine/threonine protein phosphatases, which have rather broad specificity.

15–7 Even though plants and animals independently evolved multicellularity, they use virtually all the same signaling proteins and second messengers for cell–cell communication.

Discuss the following problems.

15–8 Suppose that the circulating concentration of hormone is 10^{-10} M and the K_d for binding to its receptor is 10^{-8} M. What fraction of the receptors will have hormone bound? If a meaningful physiological response occurs when 50% of the receptors have bound a hormone molecule, how much will the concentration of hormone have to rise to elicit a response? The fraction of receptors (R) bound to hormone (H) to form a receptor–hormone complex (R–H) is $[R–H]/([R] + [R–H]) = [R–H]/[R]_{TOT} = [H]/([H] + K_d)$.

15–9 Cells communicate in ways that resemble human communication. Decide which of the following forms of human communication are analogous to autocrine, paracrine, endocrine, and synaptic signaling by cells.

A. A telephone conversation
B. Talking to people at a cocktail party
C. A radio announcement
D. Talking to yourself

15–10 Why do signaling responses that involve changes in proteins already present in the cell occur in milliseconds to seconds, whereas responses that require changes in gene expression require minutes to hours?

15–11 How is it that different cells can respond in different ways to exactly the same signaling molecule even when they have identical receptors?

15–12 Why do you suppose that phosphorylation/dephosphorylation, as opposed to allosteric binding of small molecules, for example, has evolved to play such a prominent role in switching proteins on and off in signaling pathways?

15–13 Consider a signaling pathway that proceeds through three protein kinases that are sequentially activated by phosphorylation. In one case, the kinases are held in a signaling complex by a scaffolding protein; in the other, the kinases are freely diffusible (**Figure Q15–1**). Discuss the properties of these two types of organization in terms of signal amplification, speed, and potential for cross-talk between signaling pathways.

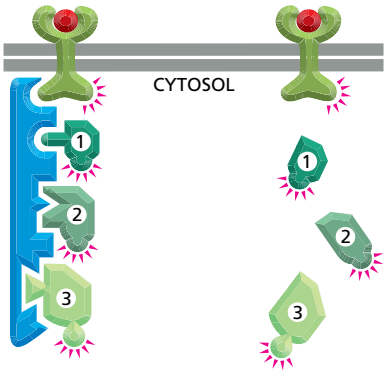


Figure Q15–1 A protein kinase cascade organized by a scaffolding protein or composed of freely diffusing components (Problem 15–13).

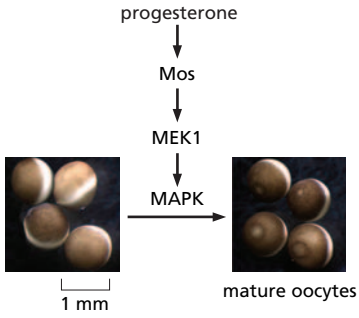


Figure Q15–2 Progesterone-induced MAP kinase activation, leading to oocyte maturation (Problem 15–15). (Courtesy of Helfrid Hochegger.)

15–14 Describe three ways in which a gradual increase in an extracellular signal can be sharpened to produce an abrupt or nearly all-or-none cellular response.

15–15 Activation ('maturation') of frog oocytes is signaled through a MAP kinase signaling module. An increase in the hormone progesterone triggers the module by stimulating the translation of Mos mRNA, which is the frog's MAP kinase kinase kinase (**Figure Q15–2**). Maturation is easy to score visually by the presence of a white spot in the middle of the brown surface of the oocyte (see Figure Q15–2). To determine the dose–response curve for progesterone-induced activation of MAP kinase, you place 16 oocytes in each of six plastic dishes and add various concentrations of progesterone. After an overnight incubation, you crush the oocytes, prepare an extract, and determine the state of MAP kinase phosphorylation (hence, activation) by SDS polyacrylamide-gel electrophoresis (**Figure Q15–3**A). This analysis shows a graded response of MAP kinase to increasing concentrations of progesterone.

Before you crushed the oocytes you noticed that not all oocytes in individual dishes had white spots. Had some oocytes undergone partial activation and not yet reached the white-spot stage? To answer this question, you repeat the experiment, but this time you analyze MAP kinase activation in individual oocytes. You are surprised to find that each oocyte has either a fully activated or a completely inactive MAP kinase (Figure Q15–3B). How can an all-or-none response in individual oocytes give rise to a graded response in the population?

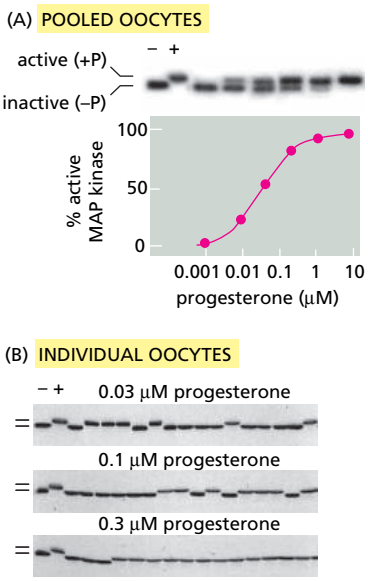


Figure Q15–3 Activation of frog oocytes (Problem 15–15). (A) Phosphorylation of MAP kinase in pooled oocytes. (B) Phosphorylation of MAP kinase in individual oocytes. MAP kinase was detected by immunoblotting using a MAP-kinase-specific antibody. The first two lanes in each gel contain nonphosphorylated, inactive MAP kinase (–) and phosphorylated, active MAP kinase (+). (From J.E. Ferrell, Jr., and E.M. Machleder, *Science* 280:895–898, 1998. With permission from AAAS.)

15–16 Propose specific types of mutations in the gene for the regulatory subunit of PKA that could lead to either a permanently active PKA or a permanently inactive PKA.

15–17 Phosphorylase kinase integrates signals from the cyclic-AMP-dependent and Ca^{2+}-dependent signaling pathways that control glycogen breakdown in liver and muscle cells (**Figure Q15–4**). Phosphorylase kinase is composed of four subunits. One is the protein kinase that catalyzes the addition of phosphate to glycogen phosphorylase to activate it for glycogen breakdown. The other three subunits are regulatory proteins that control the activity of the catalytic subunit. Two contain sites for phosphorylation by PKA, which is activated by cyclic AMP. The remaining subunit is calmodulin, which binds Ca^{2+} when the cytosolic concentration rises. The regulatory subunits control the equilibrium between the active and inactive conformations of the catalytic subunit. How does this arrangement allow phosphorylase kinase to serve its role as an integrator protein for the multiple pathways that stimulate glycogen breakdown?

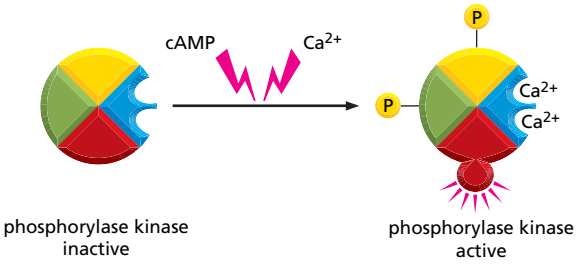


Figure Q15–4 Integration of cyclic-AMP-dependent and Ca^{2+}-dependent signaling pathways by phosphorylase kinase in liver and muscle cells (Problem 15–17).

15–18 In principle, the activated, GTP-bound form of Ras could be increased by activating a guanine-nucleotide exchange factor (GEF) or by inactivating a GTPase-activating protein (GAP). Why do you suppose that Ras-mediated signaling pathways always increase Ras–GTP by activating a GEF rather than inactivating a GAP?

15–19 The Wnt planar polarity signaling pathway normally ensures that each wing cell in *Drosophila* has a single hair. Overexpression of the *Frizzled* gene from a heat-shock promoter (hs-*Fz*) causes multiple hairs to grow from many cells (**Figure Q15–5**A). This phenotype is suppressed if hs-*Fz* is combined with a heterozygous deletion (Dsh^{Δ}) of the *Dishevelled* gene (Figure Q15–5B). Do these results allow you to order the action of Frizzled and Dishevelled in the signaling pathway? If so, what is the order? Explain your reasoning.

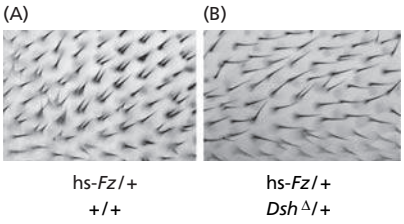


Figure Q15–5 Pattern of hair growth on wing cells in genetically different *Drosophila* (Problem 15–19). (From C.G. Winter et al., *Cell* 105:81–91, 2001. With permission from Elsevier.)

REFERENCES

General

Bradshaw RA & Dennis EA (eds) (2003) Handbook of Cell Signaling. Elsevier: St Louis.

Science's Signal Transduction Knowledge Environment (Stke): www.stke.org

General Principles of Cell Communication

Ben-Shlomo I, Yu Hsu S, Rauch R et al (2003) Signaling receptome: a genomic and evolutionary perspective of plasma membrane receptors involved in signal transduction. *Sci STKE* 187:RE9.

Bourne HR (1995) GTPases: a family of molecular switches and clocks. *Philos Trans R Soc Lond B Biol Sci* 349:283–289.

Ferrell JE Jr (2002) Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr Opin Cell Biol* 14:140–148.

Mangelsdorf DJ, Thummel C, Beato M et al (1995) The nuclear receptor superfamily: the second decade. *Cell* 83:835–839.

Murad F (2006) Shattuck Lecture. Nitric oxide and cyclic GMP in cell signaling and drug development. *N Engl J Med* 355:2003–2011.

Papin JA, Hunter T, Palsson BO & Subramaniam S (2005) Reconstruction of cellular signalling networks and analysis of their properties. *Nature Rev Mol Cell Biol* 6:99–111.

Pawson T & Scott JD (2005) Protein phosphorylation in signaling—50 years and counting. *Trends Biochem Sci* 30:286–290.

Pires-daSilva A & Sommer RJ (2003) The evolution of signalling pathways in animal development. *Nature Rev Genet* 4:39–49.

Robinson-Rechavi M, Escriva Garcia H & Laudet V (2003) The nuclear receptor superfamily. *J Cell Sci* 116:585–586.

Seet BT, Dikic I, Zhou MM & Pawson T (2006) Reading protein modifications with interaction domains. *Nature Rev Mol Cell Biol* 7:473–483.

Singla V & Reiter JF (2006) The primary cilium as the cell's antenna: signaling at a sensory organelle. *Science* 313:629–633.

Signaling Through G-protein-coupled Cell-surface Receptors and Small Intracellular Mediators

Berridge MJ (2005) Unlocking the secrets of cell signaling. *Annu Rev Physiol* 67:1–21.

Berridge MJ, Bootman MD & Roderick HL (2003) Calcium signalling: dynamics, homeostasis and remodelling. *Nature Rev Mol Cell Biol* 4:517–529.

Breer H (2003) Sense of smell: recognition and transduction of olfactory signals. *Biochem Soc Trans* 31:113–116.

Burns ME & Baylor DA (2001) Activation, deactivation, and adaptation in vertebrate photoreceptor cells. *Annu Rev Neurosci* 24:779–805.

Cooper DM (2005) Compartmentalization of adenylate cyclase and cAMP signalling. *Biochem Soc Trans* 33:1319–1322.

Hoeflich KP & Ikura M (2002) Calmodulin in action: diversity in target recognition and activation mechanisms. *Cell* 108:739–742.

Hudmon A & Schulman H (2002) Structure-function of the multifunctional Ca^{2+}/calmodulin-dependent protein kinase II. *Biochem J* 364:593–611.

Kamenetsky M, Middelhaufe S, Bank EM et al (2006) Molecular details of cAMP generation in mammalian cells: a tale of two systems. *J Mol Biol* 362:623–639.

Luttrell LM (2006) Transmembrane signaling by G protein-coupled receptors. *Methods Mol Biol* 332:3–49.

McConnachie G, Langeberg LK & Scott JD (2006) AKAP signaling complexes: getting to the heart of the matter. *Trends Mol Med* 12:317–323.

Parker PJ (2004) The ubiquitous phosphoinositides. *Biochem Soc Trans* 32:893–898.

Pierce KL, Premont RT & Lefkowitz RJ (2002) Seven-transmembrane receptors. *Nature Rev Mol Cell Biol* 3:639–650.

Reiter E & Lefkowitz RJ (2006) GRKs and beta-arrestins: roles in receptor silencing, trafficking and signaling. *Trends Endocrinol Metab* 17:159–165.

Rhee SG (2001) Regulation of phosphoinositide-specific phospholipase C. *Annu Rev Biochem* 70:281–312.

Robishaw JD & Berlot CH (2004) Translating G protein subunit diversity into functional specificity. *Curr Opin Cell Biol* 16:206–209.

Shaywitz AJ & Greenberg ME (1999) CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals. *Annu Rev Biochem* 68:821–861.

Signaling Through Enzyme-coupled Cell-surface Receptors

Baker MD, Wolanin PM & Stock JB (2006) Signal transduction in bacterial chemotaxis. *BioEssays* 28:9–22.

Dard N & Peter M (2006) Scaffold proteins in MAP kinase signaling: more than simple passive activating platforms. *BioEssays* 28:146–156.

Darnell JE Jr, Kerr IM & Stark GR (1994) Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* 264:1415–1421.

Downward J (2004) PI 3-kinase, Akt and cell survival. *Semin Cell Dev Biol* 15:177–182.

Jaffe AB & Hall A (2005) Rho GTPases: biochemistry and biology. *Annu Rev Cell Dev Biol* 21:247–269.

Massague J & Gomis RR (2006) The logic of TGFβ signaling. *FEBS Lett* 580:2811–2820.

Mitin N, Rossman KL & Der CJ (2005) Signaling interplay in Ras superfamily function. *Curr Biol* 15:R563–574.

Murai KK & Pasquale EB (2003) Eph'ective signaling: forward, reverse and crosstalk. *J Cell Sci* 116:2823–2832.

Pawson T (2004) Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell* 116:191–203.

Qi M & Elion EA (2005) MAP kinase pathways. *J Cell Sci* 118:3569–3572.

Rawlings JS, Rosler KM & Harrison DA (2004) The JAK/STAT signaling pathway. *J Cell Sci* 117:1281–1283.

Roskoski R Jr (2004) Src protein-tyrosine kinase structure and regulation. *Biochem Biophys Res Commun* 324:1155–1164.

Schlessinger J (2000) Cell signaling by receptor tyrosine kinases. *Cell* 103:211–225.

Sahin M, Greer PL, Lin MZ et al (2005) Eph-dependent tyrosine phosphorylation of ephexin1 modulates growth cone collapse. *Neuron* 46:191–204.

Schwartz MA & Madhani HD (2004) Principles of MAP kinase signaling specificity in *Saccharomyces cerevisiae*. *Annu Rev Genet* 38:725–748.

Shaw RJ & Cantley LC (2006) Ras, PI(3)K and mTOR signalling controls tumour cell growth. *Nature* 44:424–430.

Wassarman DA, Therrien M & Rubin GM (1995) The Ras signaling pathway in *Drosophila*. *Curr Opin Genet Dev* 5:44–50.

Wullschleger S, Loewith R & Hall MN (2006) TOR signaling in growth and metabolism. *Cell* 124:471–484.

Signaling Pathways Dependent on Regulated Proteolysis of Latent Gene Regulatory Proteins

Bray SJ (2006) Notch signalling: a simple pathway becomes complex. *Nature Rev Mol Cell Biol* 7:678–689.

Clevers H (2006) Wntβ-catenin signaling in development and disease. *Cell* 127:469–480.

Hoffmann A & Baltimore D (2006) Circuitry of nuclear factor κB signaling. *Immunol Rev* 210:171–186.

Huangfu D & Anderson KV (2006) Signaling from Smo to Ci/Gli: conservation and divergence of Hedgehog pathways from *Drosophila* to vertebrates. *Development* 133:3–14.

Signaling in Plants

Benavente LM & Alonso JM (2006) Molecular mechanisms of ethylene signaling in *Arabidopsis*. *Mol Biosyst* 2:165–173.

Chen M, Chory J & Fankhauser C (2004) Light signal transduction in higher plants. *Annu Rev Genet* 38:87–117.

Dievart A & Clark SE (2004) LRR-containing receptors regulating plant development and defense. *Development* 131:251–261.

Teale WD, Paponov IA & Palme K (2006) Auxin in action: signalling, transport and the control of plant growth and development. *Nature Rev Mol Cell Biol* 7:847–859.

16 The Cytoskeleton

For cells to function properly, they must organize themselves in space and interact mechanically with their environment. They have to be correctly shaped, physically robust, and properly structured internally. Many have to change their shape and move from place to place. All cells have to be able to rearrange their internal components as they grow, divide, and adapt to changing circumstances. Eucaryotic cells have developed all these spatial and mechanical functions to a very high degree, and they depend on a remarkable system of filaments called the **cytoskeleton** (Figure 16–1).

The cytoskeleton pulls the chromosomes apart at mitosis and then splits the dividing cell into two. It drives and guides the intracellular traffic of organelles, ferrying materials from one part of the cell to another. It supports the fragile plasma membrane and provides the mechanical linkages that let the cell bear stresses and strains without being ripped apart as the environment shifts and changes. It enables cells such as sperm to swim and others, such as fibroblasts and white blood cells, to crawl across surfaces. It provides the machinery in the muscle cell for contraction and in the nerve cell to extend an axon and dendrites. It guides the growth of the plant cell wall and controls the amazing diversity of eucaryotic cell shapes.

The cytoskeleton's varied functions depend on the behavior of three families of protein molecules, which assemble to form three main types of filaments. Each type of filament has distinct mechanical properties, dynamics, and biological roles, but all three share certain fundamental principles. These principles provide the basis for a general understanding of how the cytoskeleton works and how the different elements cooperate. Just as we require our ligaments, bones, and muscles to work together, so all three cytoskeletal filament systems must normally function collectively to give a cell its strength, its shape, and its ability to move.

In this chapter, we begin by describing the three main types of filaments, the basic principles underlying their assembly and disassembly, and their individual peculiarities. We then describe how other proteins interact with the three main filament systems, enabling the cell to establish and maintain internal order, to shape and remodel its surface, to move organelles in a directed manner from one place to another, and—when appropriate—to move itself to new locations.

In This Chapter

THE SELF-ASSEMBLY AND DYNAMIC STRUCTURE OF CYTOSKELETAL FILAMENTS 965

HOW CELLS REGULATE THEIR CYTOSKELETAL FILAMENTS 992

MOLECULAR MOTORS 1010

THE CYTOSKELETON AND CELL BEHAVIOR 1025

THE SELF-ASSEMBLY AND DYNAMIC STRUCTURE OF CYTOSKELETAL FILAMENTS

Most animal cells have three types of cytoskeletal filaments that are responsible for the cells' spatial organization and mechanical properties. *Intermediate filaments* provide mechanical strength. *Microtubules* determine the positions of membrane-enclosed organelles and direct intracellular transport. *Actin filaments* determine the shape of the cell's surface and are necessary for whole-cell locomotion. But these cytoskeletal filaments would be ineffective without the hundreds of accessory proteins that link the filaments to other cell components, as well as to each other. This large set of *accessory proteins* is essential for the

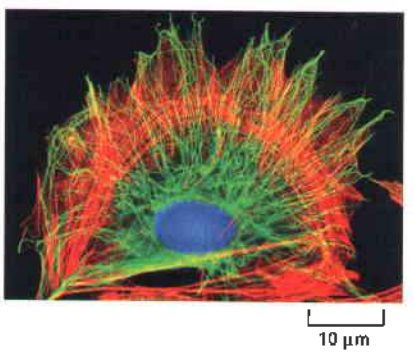


Figure 16–1 The cytoskeleton. A cell in culture has been fixed and labeled to show two of its major cytoskeletal systems, the microtubules (*green*) and the actin filaments (*red*). The DNA in the nucleus is labeled in *blue*. (Courtesy of Albert Tousson.)

controlled assembly of the cytoskeletal filaments in particular locations, and it includes the *motor proteins*, remarkable molecular machines that convert the energy of ATP hydrolysis into mechanical force that can either move organelles along the filaments or move the filaments themselves.

In this section, we discuss the properties of the proteins that make up the filaments of the cytoskeleton. We focus on their ability to form intrinsically polarized and self-organized structures. We shall see that, because of the remarkable mechanisms that cause cytoskeletal filaments to be dynamic, the cell is able to respond rapidly to any eventuality that it may face.

Cytoskeletal Filaments Are Dynamic and Adaptable

Cytoskeletal systems are dynamic and adaptable, organized more like ant trails than interstate highways. A single trail of ants may persist for many hours, extending from the ant nest to a delectable picnic site, but the individual ants within the trail are anything but static. If the ant scouts find a new and better source of food, or if the picnickers clean up and leave, the dynamic structure rearranges itself with astonishing rapidity to deal with the new situation. In a similar way, large-scale cytoskeletal structures can change or persist, according to need, lasting for lengths of time ranging from less than a minute up to the cell's lifetime. But the individual macromolecular components that make up these structures are in a constant state of flux. Thus, like the alteration of an ant trail, a structural rearrangement in a cell requires little extra energy when conditions change.

Regulation of the dynamic behavior and assembly of the cytoskeletal filaments allows eucaryotic cells to build an enormous range of structures from the three basic filament systems. The micrographs in **Panel 16–1** reveal some of these structures. Microtubules, which are frequently found in a star-like cytoplasmic array emanating from the center of an interphase cell, can quickly rearrange themselves to form a bipolar *mitotic spindle* during cell division. They can also form motile whips called *cilia* and *flagella* on the surface of the cell, or tightly aligned bundles that serve as tracks for the transport of materials down long neuronal axons. In plant cells, organized arrays of microtubules help to direct the pattern of cell wall synthesis.

Actin filaments underlie the plasma membrane of animal cells, providing strength and shape to its thin lipid bilayer. They also form many types of cell-surface projections. Some of these are dynamic structures, such as the *lamellipodia* and *filopodia* that cells use to explore territory and pull themselves around. The actin-based *contractile ring* assembles transiently to divide cells in two; more stable arrays allow cells to brace themselves against an underlying substratum and enable muscle to contract. The regular bundles of *stereocilia* on the surface of hair cells in the inner ear contain stable bundles of actin filaments that tilt as rigid rods in response to sound, and similarly organized *microvilli* on the surface of intestinal epithelial cells vastly increase the apical cell surface area to enhance nutrient absorption.

Intermediate filaments line the inner face of the nuclear envelope, forming a protective cage for the cell's DNA; in the cytosol, they are twisted into strong cables that can hold epithelial cell sheets together or help nerve cells to extend long and robust axons, and they allow us to form tough appendages such as hair and fingernails.

An important and dramatic example of rapid reorganization of the cytoskeleton occurs during cell division, as shown in **Figure 16–2** for a fibroblast growing in a tissue culture dish. After the chromosomes have replicated, the interphase microtubule array that spans throughout the cytoplasm is reconfigured into the bipolar *mitotic spindle*, which serves the critical function of accurately segregating the two copies of each replicated chromosome into two separate daughter nuclei. At the same time, the specialized actin structures that enable the fibroblast to crawl across the surface of the dish disassemble so that the cell stops moving, rounds up, and assumes a more spherical shape. Actin and its associated motor protein myosin then form a belt around the middle of the cell, the *contractile ring*, which constricts like a tiny muscle to pinch the cell

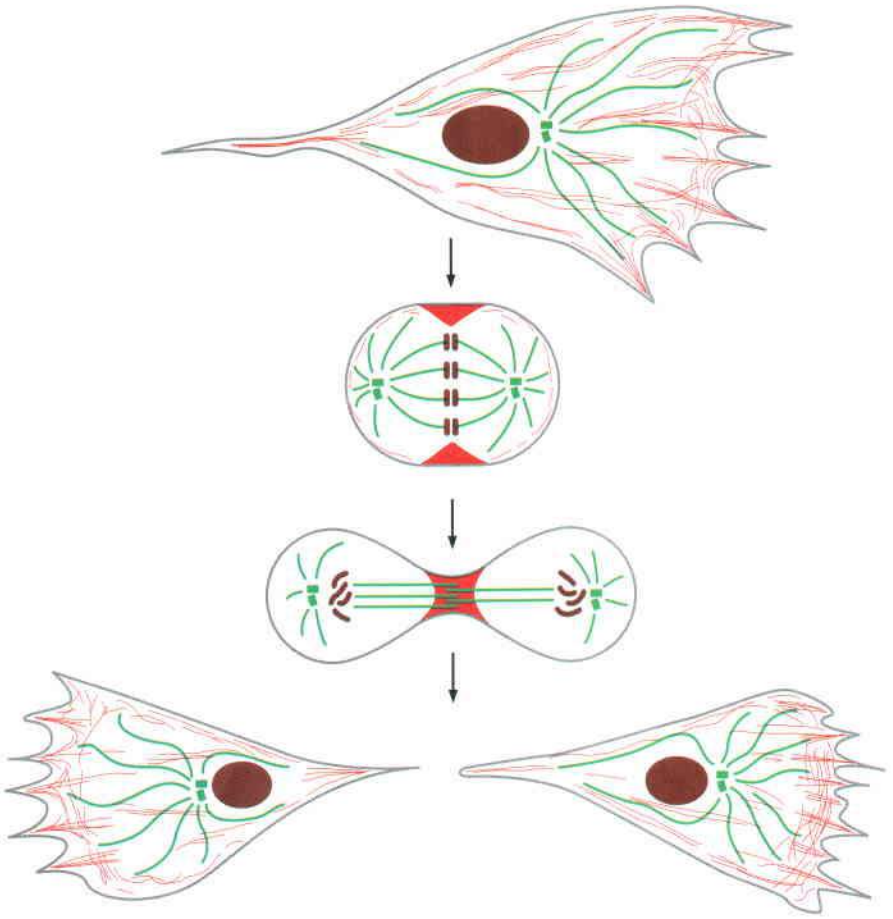


Figure 16–2 Rapid changes in cytoskeletal organization associated with cell division. The crawling fibroblast drawn here has a polarized, dynamic actin cytoskeleton (shown in *red*) that assembles to push its leading edge toward the right. The polarization of the actin cytoskeleton is assisted by the microtubule cytoskeleton (shown in *green*), consisting of long microtubules that emanate from a single microtubule organizing center located in front of the nucleus. When the cell divides, the polarized microtubule array rearranges to form a bipolar mitotic spindle, which is responsible for aligning and then separating the duplicated chromosomes (*brown*). The actin filaments form a contractile ring at the center of the cell that pinches the cell in two after the chromosomes have separated. After cell division is complete, the two daughter cells reorganize both the microtubule and actin cytoskeletons into smaller versions of those that were present in the mother cell, enabling them to crawl their separate ways.

in two. When division is complete, the cytoskeletons of the two daughter fibroblasts reassemble into their interphase structures to convert the two rounded up daughter cells into smaller versions of the flattened, crawling mother cell. In a fibroblast, this sequence of events takes about an hour; in some cases, such as the early nuclear divisions in a *Drosophila* embryo, the actin and microtubule cytoskeletons can completely rearrange themselves within less than five minutes (Figure 16–3).

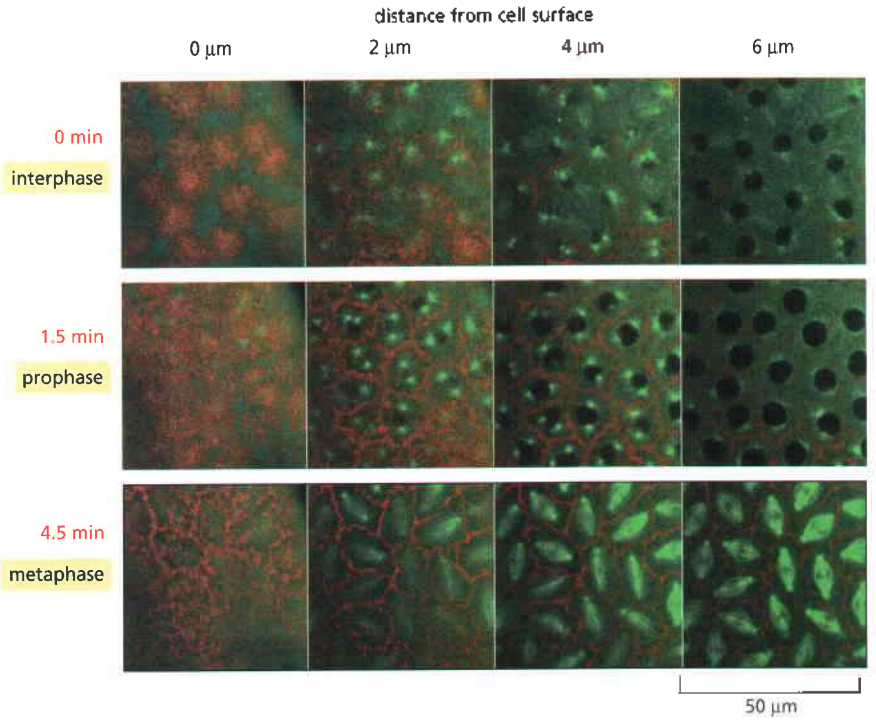
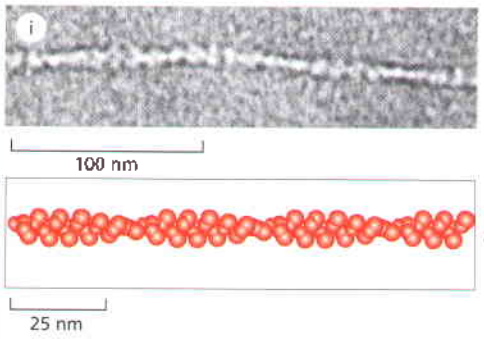
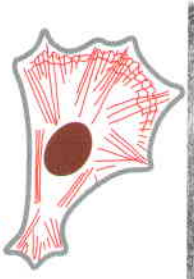
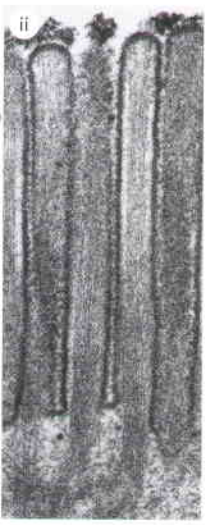


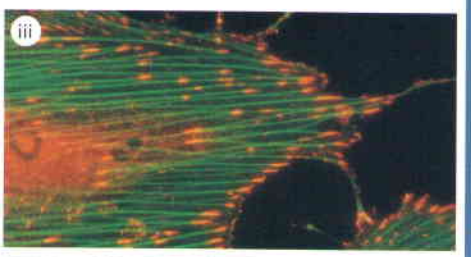
Figure 16–3 Rapid changes in cytoskeletal organization observed during the development of a *Drosophila* early embryo. <TTCT> In this giant multinuclear cell, the early nuclear divisions occur every 10 minutes or so in a common cytoplasm. The rapid rearrangements of the actin filaments *(red)* and microtubules *(green)*, seen here in a living embryo, are required to separate the chromosomes at mitosis, while keeping each nucleus from colliding with its neighbors. (Courtesy of William Sullivan.)

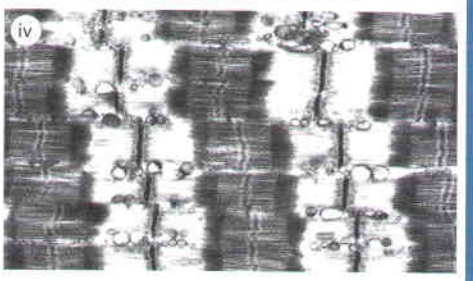
ACTIN FILAMENTS







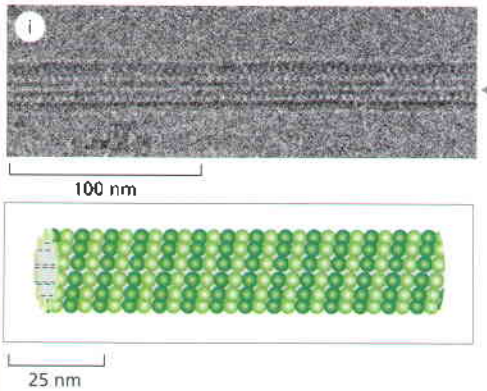


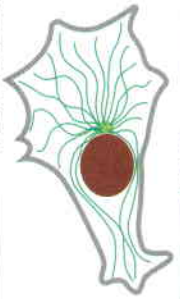


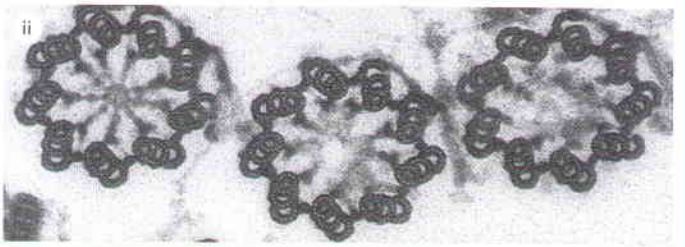
Actin filaments (also known as *microfilaments*) are two-stranded helical polymers of the protein actin. They appear as flexible structures, with a diameter of 5–9 nm, and they are organized into a variety of linear bundles, two-dimensional networks, and three-dimensional gels. Although actin filaments are dispersed throughout the cell, they are most highly concentrated in the *cortex*, just beneath the plasma membrane.

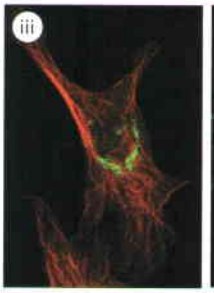
Micrographs courtesy of Roger Craig (i and iv); P.T. Matsudaira and D.R. Burgess (ii); Keith Burridge (iii).

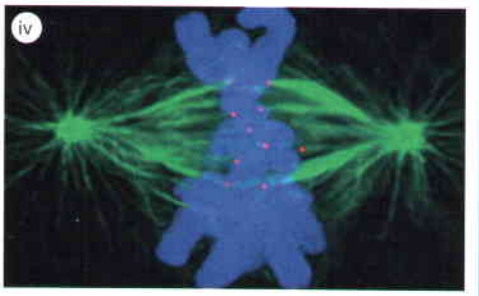
MICROTUBULES







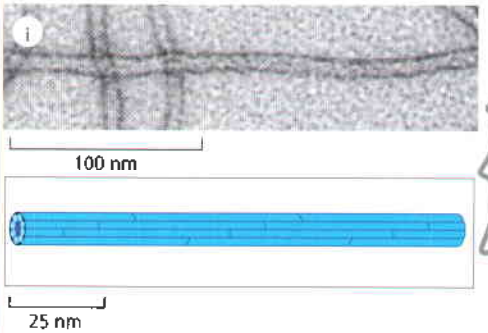




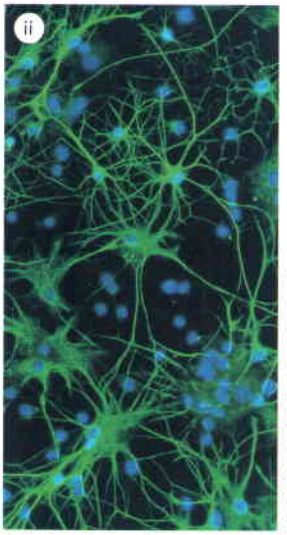
Microtubules are long, hollow cylinders made of the protein tubulin. With an outer diameter of 25 nm, they are much more rigid than actin filaments. Microtubules are long and straight and typically have one end attached to a single microtubule-organizing center (MTOC) called a *centrosome*.

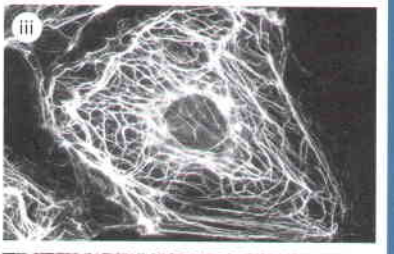
Micrographs courtesy of Richard Wade (i); D.T. Woodrow and R.W. Linck (ii); David Shima (iii); A. Desai (iv).

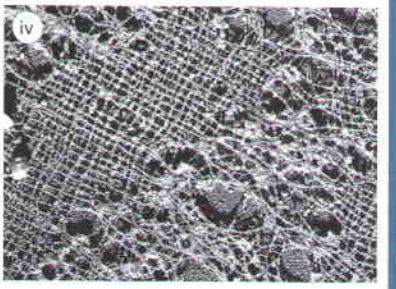
INTERMEDIATE FILAMENTS











Intermediate filaments are ropelike fibers with a diameter of around 10 nm; they are made of intermediate filament proteins, which constitute a large and heterogeneous family. One type of intermediate filament forms a meshwork called the nuclear lamina just beneath the inner nuclear membrane. Other types extend across the cytoplasm, giving cells mechanical strength. In an epithelial tissue, they span the cytoplasm from one cell-cell junction to another, thereby strengthening the entire epithelium.

Micrographs courtesy of Roy Quinlan (i); Nancy L. Kedersha (ii); Mary Osborn (iii); Ueli Aebi (iv).

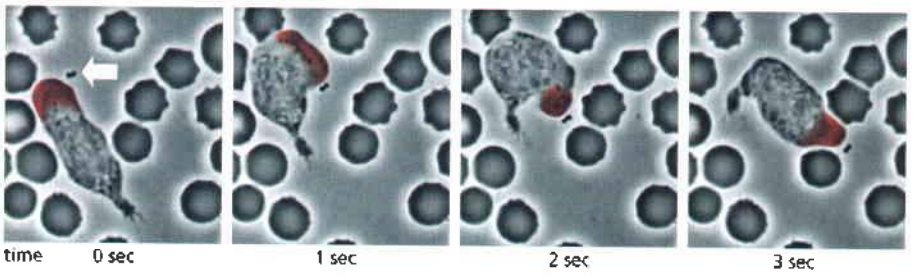


Figure 16–4 A neutrophil in pursuit of bacteria. <TGTA> In this preparation of human blood, a clump of bacteria (*white arrow*) is about to be captured by a neutrophil. As the bacteria move, the neutrophil quickly reassembles the dense actin network at the leading edge (*red*) to push toward the location of the bacteria. Rapid disassembly and reassembly of the actin cytoskeleton in this cell enables it to change its orientation and direction of movement within a few seconds. (From a video recorded by David Rogers.)

Many cells require rapid cytoskeletal rearrangements for their normal functioning during interphase as well. For example, the *neutrophil*, a type of white blood cell, chases and engulfs bacterial and fungal cells that accidentally gain access to the normally sterile parts of the body, as through cuts in the skin. Like most crawling cells, neutrophils advance by extending a protrusive structure at the leading edge filled with newly polymerized actin filaments. When the elusive bacterial prey moves in a different direction, the neutrophil is poised to reorganize its polarized protrusive structures within seconds (**Figure 16–4**). Both of these kinds of rapid cytoskeletal rearrangements will be discussed in more detail in the final section of this chapter.

The Cytoskeleton Can Also Form Stable Structures

In cells that have achieved a stable, differentiated morphology such as mature neurons or epithelial cells, the dynamic elements of the cytoskeleton must also provide stable, large-scale structures for cellular organization. On specialized epithelial cells that line tissues such as the intestine and the lung, cytoskeletal-based cell surface protrusions including microvilli and cilia are able to maintain a constant location, length, and diameter over the entire lifetime of the cell. For the actin bundles at the cores of microvilli on intestinal epithelial cells this is only a few days. But the actin bundles at the cores of stereocilia on the hair cells of the inner ear must maintain their stable organization for the entire lifetime of the animal, since these cells do not turn over. Nonetheless, the individual actin filaments remain strikingly dynamic and are continuously remodeled and replaced on average every 48 hours, even within these stable cell surface structures that persist for decades.

Besides forming stable specialized cell surface protrusions, the cytoskeleton is also responsible for large-scale cellular polarity, enabling cells to tell the difference between top and bottom, or front and back. The large-scale polarity information encoded by the organization of the cytoskeleton must also often be maintained over the lifetime of the cell. Polarized epithelial cells such as those found in the lining of the intestine, for example, use organized arrays of microtubules, actin filaments, and intermediate filaments to maintain the critical functional differences between the *apical surface* that absorbs nutrients from the lumen of the intestine where food passes by to the *basolateral surface* where the cells transfer nutrients through the plasma membrane to the bloodstream. They also must maintain strong adhesive contacts with one another to enable this single layer of cells to serve as an effective physical barrier (**Figure 16–5**).

Even small, morphologically simple cells such as the budding yeast *Saccharomyces cerevisiae* need stable large-scale polarity. The most notable feature of the structure of these cells is the marked asymmetry, evident in the way they divide by budding to create a small daughter cell and a large mother cell. This asymmetry derives from the polar orientation of the cell's actin cytoskeleton. There are two types of actin filament assemblies in these cells: actin cables (long bundles of actin filaments) and actin patches (small assemblies of filaments associated with the cell cortex, marking sites of actin-driven endocytosis). Proliferating budding yeast cells must be highly polarized to allow the cell to grow a bud from a single site on the cell surface, as opposed to simply growing uniformly larger. In this process, the actin patches become highly concentrated at the growing tip of the bud, with the actin cables aligned and pointing toward

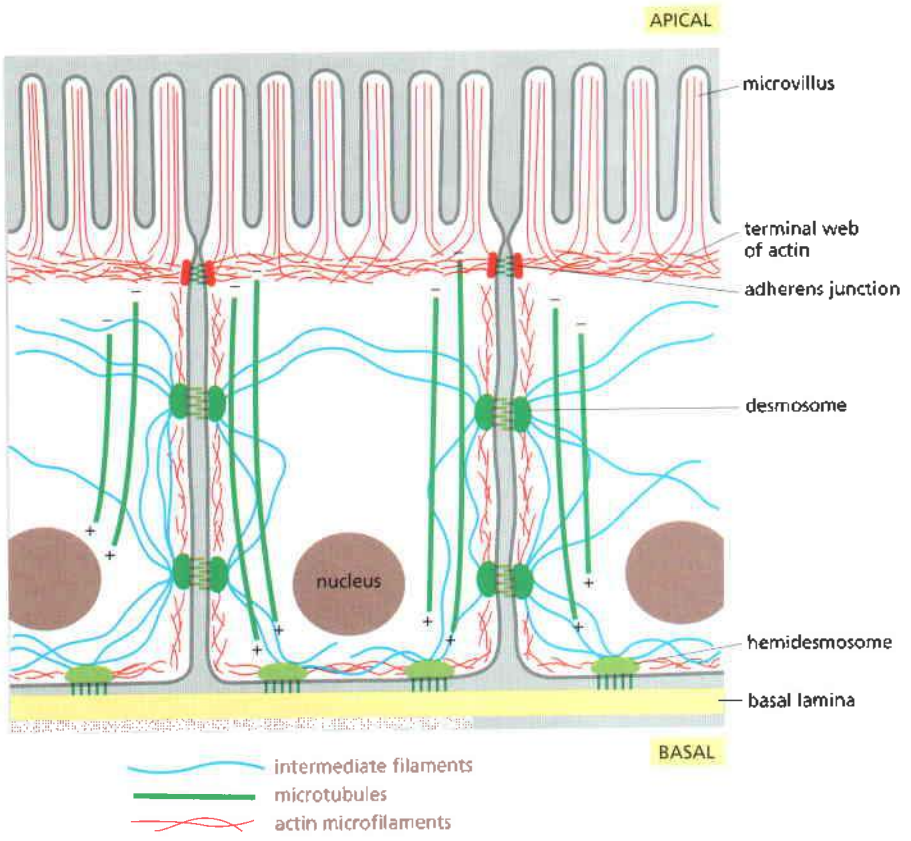


Figure 16–5 Organization of the cytoskeleton in polarized epithelial cells. All the components of the cytoskeleton cooperate to produce the characteristic shapes of specialized cells including the epithelial cells that line the small intestine. At the apical (upper) surface, facing the intestinal lumen, bundled actin filaments (*red*) form microvilli that increase the cell surface area available for absorbing nutrients from food. Just below the microvilli, a circumferential band of actin filaments contributes to forming cell-cell junctions that prevent the contents of the intestinal lumen from leaking into the body. Intermediate filaments (*blue*) are anchored to other kinds of adhesive structures including desmosomes and hemidesmosomes that connect the epithelial cells into a sturdy sheet and attach them to the underlying extracellular matrix on the basal side of the cell; these important adhesive structures will be discussed in Chapter 19. Microtubules (*green*) run vertically from the top of the cell to the bottom, and provide a global coordinate system that enables the cell to direct newly synthesized components to their proper locations.

them. This actin organization directs the secretion of new cell wall and other materials to the site of budding (**Figure 16–6**). The polarized organization of the actin structures in turn influences the orientation of the mitotic spindle, so that a complete set of replicated chromosomes can be delivered into the daughter cell at the end of the cell division process.

Each Type of Cytoskeletal Filament Is Constructed from Smaller Protein Subunits

Cytoskeletal structures frequently reach all the way from one end of the cell to the other, spanning tens or even hundreds of micrometers. Yet the individual protein molecules of the cytoskeleton are generally only a few nanometers in size. The cell builds the large structures by the repetitive assembly of large numbers of the small subunits, like building a skyscraper out of bricks. Because these subunits are small, they can diffuse rapidly within the cytoplasm, whereas the assembled filaments cannot. In this way, cells can undergo rapid

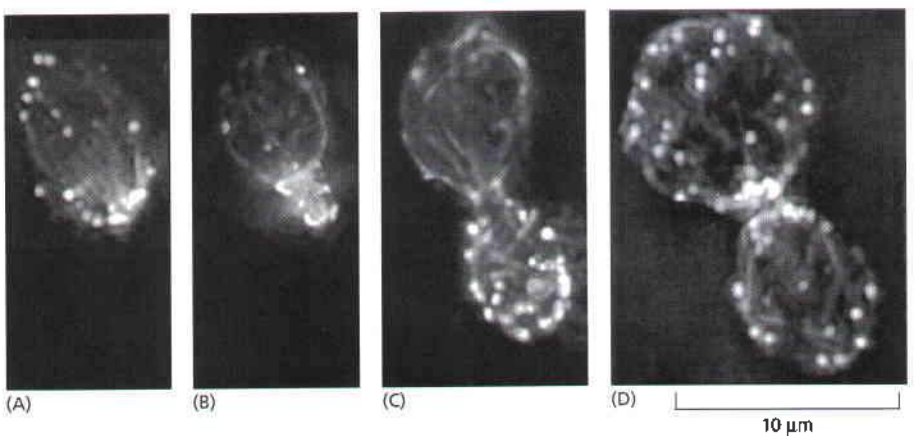


Figure 16–6 Polarity of actin patches and cables throughout the yeast cell cycle. Filamentous actin structures in the yeast cell, labeled here with fluorescent phalloidin, include actin patches (round bright spots) and actin cables (extended lines). (A) In a mother cell, before the formation of the bud, most of the patches become clustered at one end. The cables are lined up and point toward the cluster of patches, which is the site where the bud will emerge. (B) As the small bud grows, most patches remain within it. Cables in the mother cell continue to point toward this site of new cell wall growth. (C) Patches are almost uniformly distributed over the surface of a full-sized bud. Cables in the mother cell remain polarized. (D) Immediately after cell division, mother and daughter cells form new patches, which are concentrated near the division site, although both cells have randomly oriented cables. (From T.S. Karpova et al., *J. Cell Biol.* 142:1501–1517, 1998. With permission from The Rockefeller University Press.)

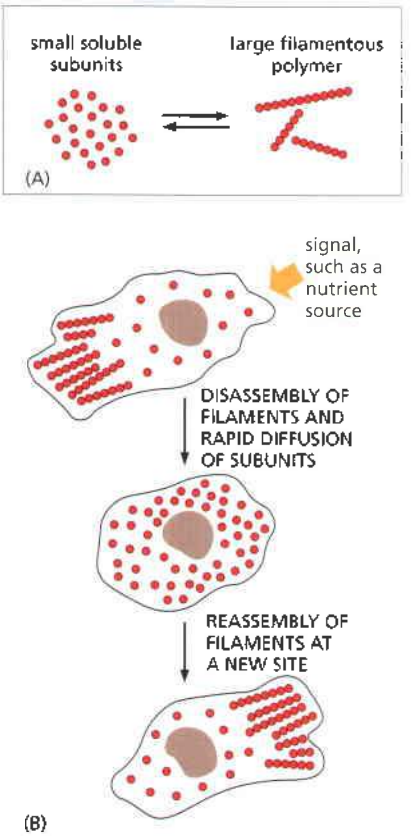


Figure 16–7 The cytoskeleton during changes in cell shape. The formation of protein filaments from much smaller protein subunits allows regulated filament assembly and disassembly to reshape the cytoskeleton. (A) Filament formation from a small protein. (B) Rapid reorganization of the cytoskeleton in a cell in response to an external signal.

structural reorganizations, disassembling filaments at one site and reassembling them at another site far away (**Figure 16–7**).

Intermediate filaments are made up of smaller subunits that are themselves elongated and fibrous, whereas actin filaments and microtubules are made of subunits that are compact and globular—*actin subunits* for actin filaments, *tubulin subunits* for microtubules. All three types of cytoskeletal filaments form as helical assemblies of subunits (see Figure 3–26) that self-associate, using a combination of end-to-end and side-to-side protein contacts. Differences in the structures of the subunits and the strengths of the attractive forces between them produce critical differences in the stability and mechanical properties of each type of filament.

Covalent linkages between their subunits hold together many biological polymers—including DNA, RNA, and proteins. In contrast, weak noncovalent interactions hold together the three types of cytoskeletal "polymers". Consequently, their assembly and disassembly can occur rapidly, without covalent bonds being formed or broken.

Within the cell, hundreds of different cytoskeleton-associated accessory proteins regulate the spatial distribution and the dynamic behavior of the filaments, converting information received through signaling pathways into cytoskeletal action. These accessory proteins bind to the filaments or their subunits to determine the sites of assembly of new filaments, to regulate the partitioning of polymer proteins between filament and subunit forms, to change the kinetics of filament assembly and disassembly, to harness energy to generate force, and to link filaments to one another or to other cell structures such as organelles and the plasma membrane. In these processes, the accessory proteins bring cytoskeletal structure under the control of extracellular and intracellular signals, including those that trigger the dramatic transformations of the cytoskeleton that occur during each cell cycle. Acting together, the accessory proteins enable a eucaryotic cell to maintain a highly organized but flexible internal structure and, in many cases, to move.

Filaments Formed from Multiple Protofilaments Have Advantageous Properties

In general, we can view the linking of protein subunits together to form a filament as a simple association reaction. A free subunit binds to the end of a filament that contains n subunits to generate a filament of length $n + 1$. The addition of each subunit to the end of the polymer creates a new end to which yet another subunit can bind. However, the robust cytoskeletal filaments in living cells are not built by simply stringing subunits together in a single straight file. A thousand tubulin monomers, for example, lined up end to end, would span the diameter of a small eucaryotic cell, but a filament formed in this way would lack the strength to avoid breakage by ambient thermal energy, unless each subunit in the filament was bound extremely tightly to its two neighbors. Such tight binding would limit the rate at which the filaments could disassemble, making the cytoskeleton a static and less useful structure.

Cytoskeletal polymers combine strength with adaptability because they are built out of multiple **protofilaments**—long linear strings of subunits joined end to end—that associate with one another laterally. Typically, the protofilaments twist around one another in a helical lattice. The addition or loss of a subunit at the end of one protofilament makes or breaks one set of longitudinal bonds and either one or two sets of lateral bonds. In contrast, breakage of the composite filament in the middle requires breaking sets of longitudinal bonds in several

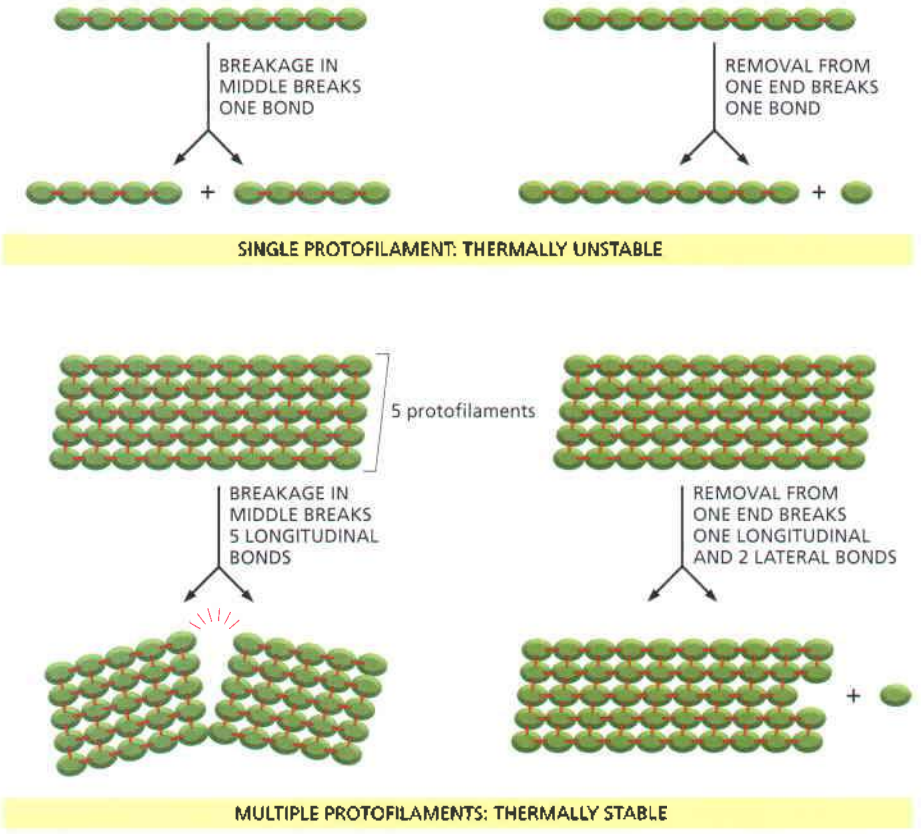


Figure 16–8 The thermal stability of cytoskeletal filaments with dynamic ends. Formation of a cytoskeletal filament from more than one protofilament allows the ends to be dynamic, while the filaments themselves are resistant to thermal breakage. In this hypothetical example, the stable filament is formed from five protofilaments. The bonds holding the subunits together in the filaments are shown in *red*.

protofilaments all at the same time (Figure 16–8). The large energy difference between these two processes allows most cytoskeletal filaments to resist thermal breakage, while leaving the filament ends as dynamic structures at which addition and loss of subunits can occur rapidly.

As with other specific protein–protein interactions, many hydrophobic interactions and weak noncovalent bonds hold the subunits in a cytoskeletal filament together (see Figure 3–4). The locations and types of subunit–subunit contacts differ for the different cytoskeletal filaments. Intermediate filaments, for example, assemble by forming strong lateral contacts between α-helical coiled coils, which extend over most of the length of each elongated fibrous subunit. Because the individual subunits are staggered in the filament, intermediate filaments tolerate stretching and bending, forming strong rope-like structures (Figure 16–9). Microtubules, by contrast, are built from globular subunits held together primarily by longitudinal bonds, and the lateral bonds holding the 13 protofilaments together are comparatively weak. For this reason, microtubules break much more easily when they are bent than do intermediate filaments.

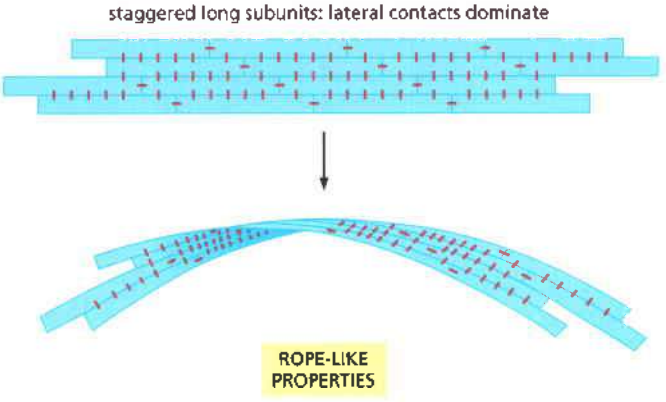


Figure 16–9 A strong filament formed from elongated fibrous subunits with strong lateral contacts. Intermediate filaments are formed in this way and are consequently especially resistant to stretching forces, although they are easily bent.

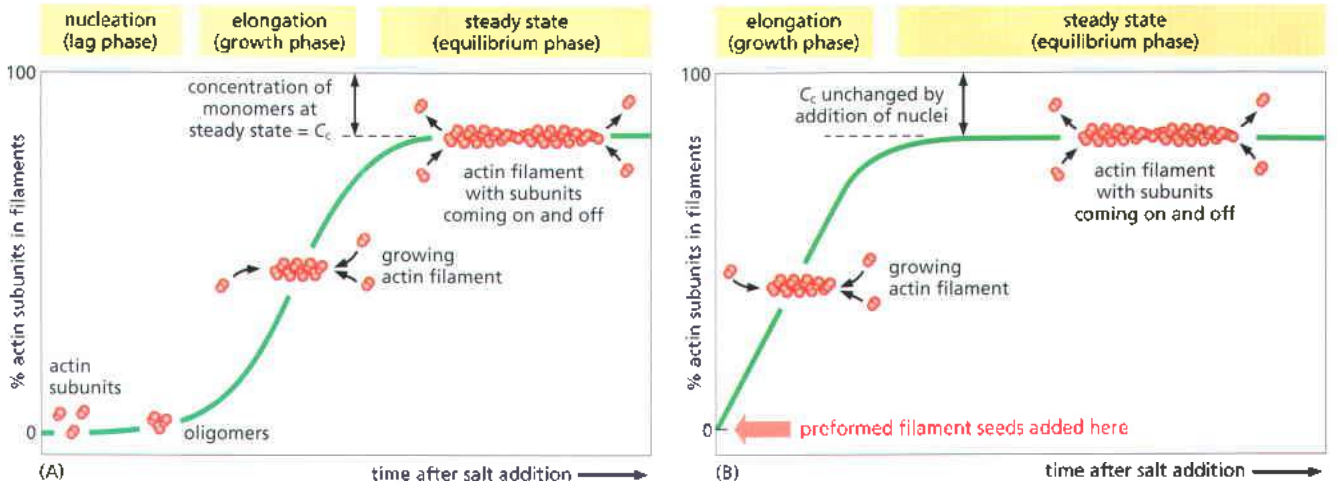


Figure 16–10 **The time course of actin polymerization in a test tube.** (A) Polymerization is begun by raising the salt concentration in a solution of pure actin subunits. (B) Polymerization is begun in the same way, but with preformed fragments of actin filaments present to act as nuclei for filament growth. As indicated, the % free subunits reflects the critical concentration (C_c), the point at which there is no net change in polymer.

Nucleation Is the Rate-Limiting Step in the Formation of a Cytoskeletal Polymer

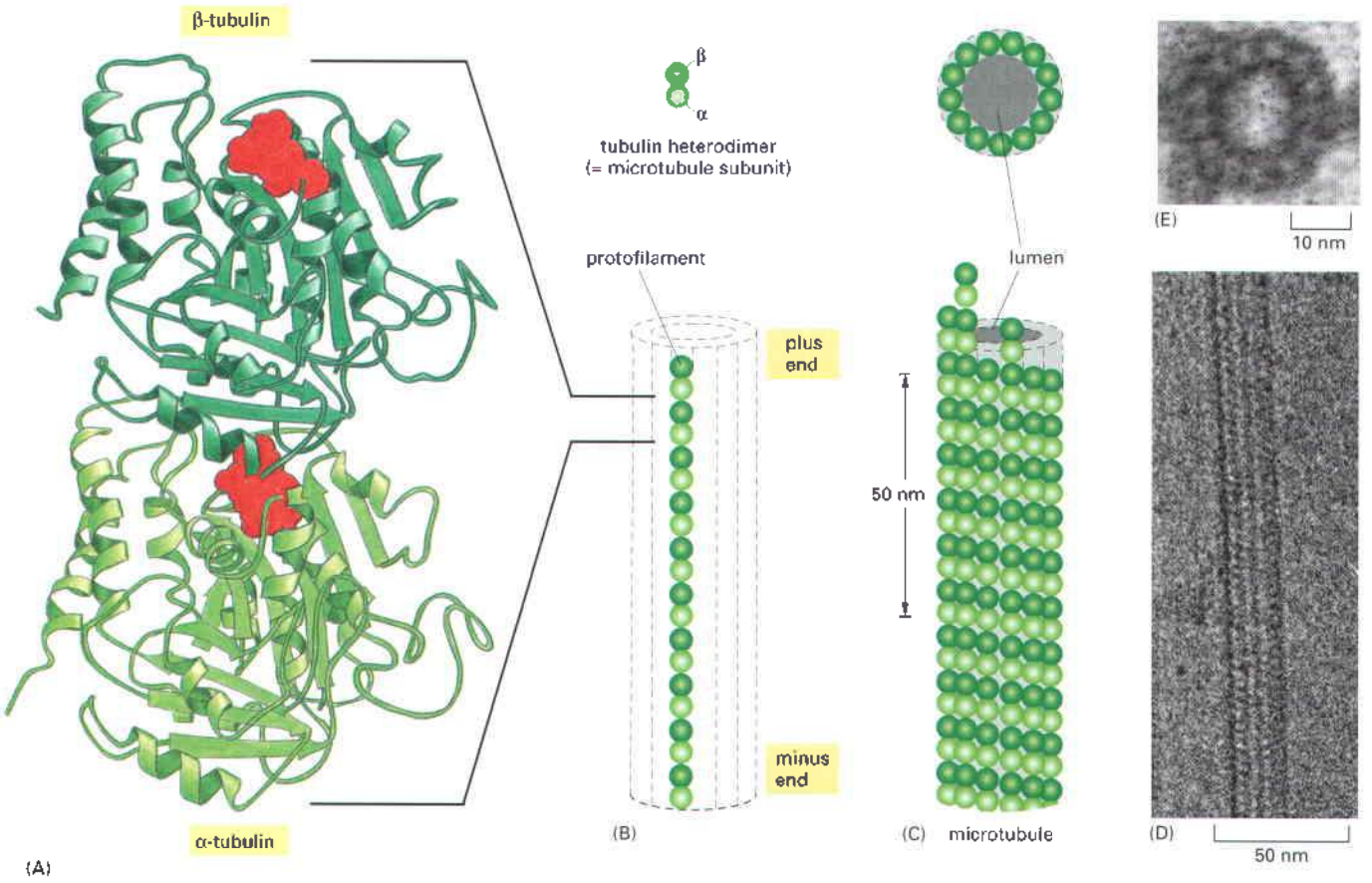
There is an important additional consequence of the multiple-protofilament organization of cytoskeletal polymers. Short oligomers composed of a few subunits can assemble spontaneously, but they are unstable and disassemble readily because each monomer is bonded only to a few other monomers. For a new large filament to form, subunits must assemble into an initial aggregate, or nucleus, that is stabilized by many subunit–subunit contacts and can then elongate rapidly by addition of more subunits. The initial process of nucleus assembly is called filament *nucleation*, and it can take quite a long time, depending on how many subunits must come together to form the nucleus.

The instability of smaller aggregates creates a kinetic barrier to nucleation, which is easily observed in a solution of pure actin or tubulin—the subunits of actin filaments and microtubules, respectively. When polymerization is initiated in a test tube containing a solution of pure individual subunits (by raising the temperature or raising the salt concentration), there is an initial lag phase, during which no filaments are observed. During this lag phase, however, a few of the small unstable aggregates succeed in making the transition to the more stable filament form, so that the lag phase is followed by a phase of rapid filament elongation, during which subunits add quickly onto the ends of the nucleated filaments (Figure 16–10A). Finally, the system approaches a steady state at which the rate of addition of new subunits to the filament ends exactly balances the rate of subunit dissociation from the ends. The concentration of free subunits left in solution at this point is called the *critical concentration*, C_c. As explained in Panel 16–2 (pp. 978–979), the value of the critical concentration is equal to the rate constant for subunit loss divided by the rate constant for subunit addition—that is, $C_c = k_{off} / k_{on}$.

The lag phase in filament growth is eliminated if preexisting seeds (such as filament fragments that have been chemically cross-linked) are added to the solution at the beginning of the polymerization reaction (Figure 16–10B). The cell takes great advantage of this nucleation requirement: it uses special proteins to catalyze filament nucleation at specific sites, thereby determining the location at which new cytoskeletal filaments are assembled. Indeed, the regulation of filament nucleation is a primary way for cells to control their shape and their movement.

The Tubulin and Actin Subunits Assemble Head-to-Tail to Create Polar Filaments

Microtubules are formed from protein subunits of **tubulin**. The tubulin subunit is itself a heterodimer formed from two closely related globular proteins called *α-tubulin* and *β-tubulin*, tightly bound together by noncovalent bonds (Figure



16–11). These two tubulin proteins are found only in this heterodimer. Each α or β monomer has a binding site for one molecule of GTP. The GTP that is bound to the α-tubulin monomer is physically trapped at the dimer interface and is never hydrolyzed or exchanged; it can therefore be considered to be an integral part of the tubulin heterodimer structure. The nucleotide on the β-tubulin, in contrast, may be in either the GTP or the GDP form, and it is exchangeable. As we shall see, the hydrolysis of GTP at this site to produce GDP has an important effect on microtubule dynamics.

A microtubule is a hollow cylindrical structure built from 13 parallel protofilaments, each composed of alternating α-tubulin and β-tubulin molecules. When the tubulin heterodimers assemble to form the hollow cylindrical microtubule, they generate two new types of protein–protein contacts. Along the longitudinal axis of the microtubule, the "top" of one β-tubulin molecule forms an interface with the "bottom" of the α-tubulin molecule in the adjacent heterodimer. This interface is very similar to the interface holding the α and β monomers together in the dimer subunit, and the binding energy is strong. Perpendicular to these interactions, neighboring protofilaments form lateral contacts. In this dimension, the main lateral contacts are between monomers of the same type (α–α and β–β). Together, the longitudinal and lateral contacts are repeated in the regular helical lattice of the microtubule. Because multiple contacts within the lattice hold most of the subunits in a microtubule in place, the addition and loss of subunits occurs almost exclusively at the microtubule ends (see Figure 16–8). These multiple contacts among subunits make microtubules stiff and difficult to bend. The stiffness of a filament can be characterized by its *persistence length*, a property of the filament describing how long it must be before random thermal fluctuations are likely to cause it to bend. The persistence length of a microtubule is several millimeters, making microtubules the stiffest and straightest structural elements found in most animal cells.

The subunits in each protofilament in a microtubule all point in the same direction, and the protofilaments themselves are aligned in parallel (in Figure 16–11, for example, the α-tubulin is down and the β-tubulin up in each hetero-

Figure 16–11 The structure of a microtubule and its subunit. (A) The subunit of each protofilament is a tubulin heterodimer, formed from a very tightly linked pair of α- and β-tubulin monomers. The GTP molecule in the α-tubulin monomer is so tightly bound that it can be considered an integral part of the protein. The GTP molecule in the β-tubulin monomer, however, is less tightly bound and has an important role in filament dynamics. Both nucleotides are shown in *red*. (B) One tubulin subunit (α–β heterodimer) and one protofilament are shown schematically. Each protofilament consists of many adjacent subunits with the same orientation. (C) The microtubule is a stiff hollow tube formed from 13 protofilaments aligned in parallel. (D) A short segment of a microtubule viewed in an electron microscope. (E) Electron micrograph of a cross section of a microtubule showing a ring of 13 distinct protofilaments. (D, courtesy of Richard Wade; E, courtesy of Richard Linck.)

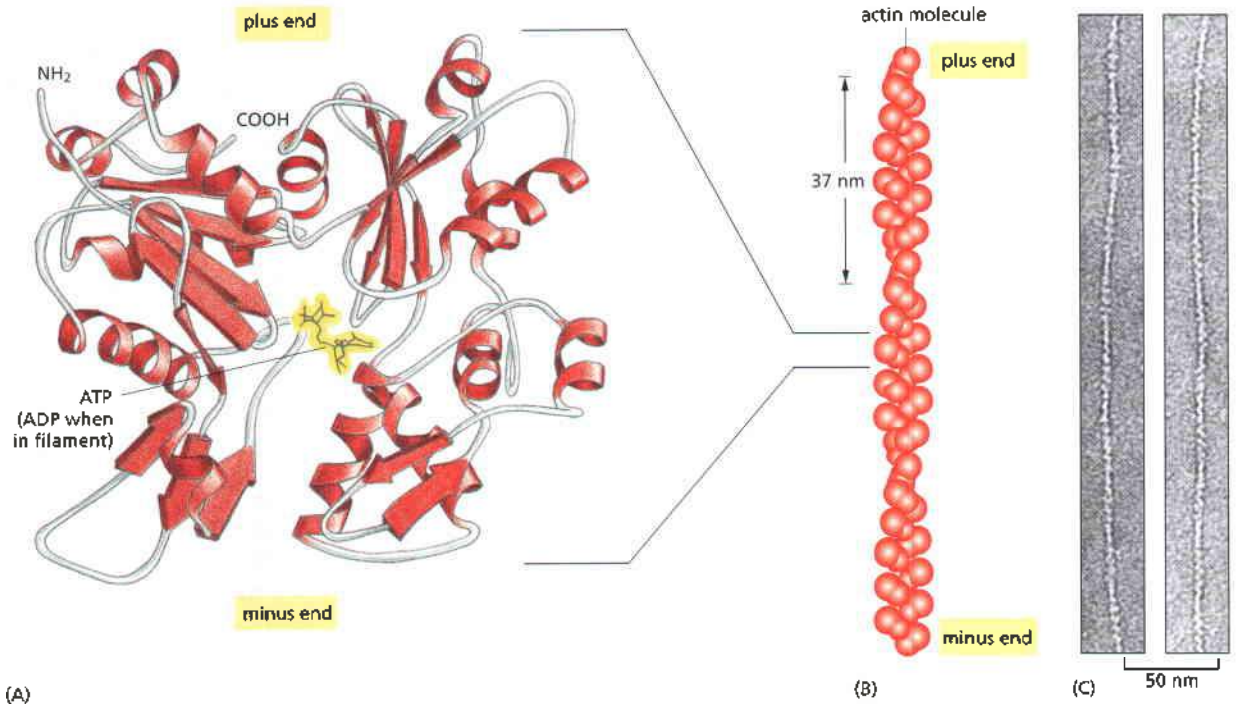


Figure 16–12 The structures of an actin monomer and actin filament. (A) The actin monomer has a nucleotide (either ATP or ADP) bound in a deep cleft in the center of the molecule. (B) Arrangement of monomers in a filament. Although the filament is often described as a single helix of monomers, it can also be thought of as consisting of two protofilaments, held together by lateral contacts, which wind around each other as two parallel strands of a helix, with a twist repeating every 37 nm. All the subunits within the filament have the same orientation. (C) Electron micrographs of negatively stained actin filaments. (C, courtesy of Roger Craig.)

dimer). Therefore, the microtubule itself has a distinct structural polarity, with α-tubulins exposed at one end and β-tubulins exposed at the other end.

The actin subunit is a single globular polypeptide chain and is thus a monomer rather than a dimer. Like tubulin, each actin subunit has a binding site for a nucleotide, but for actin the nucleotide is ATP (or ADP) rather than GTP (or GDP) (**Figure 16–12**). As for tubulin, the actin subunits assemble head-to-tail to generate filaments with a distinct structural polarity. The actin filament can be considered to consist of two parallel protofilaments that twist around each other in a right-handed helix. Actin filaments are relatively flexible and easily bent compared with the hollow cylindrical microtubules, with a persistence length of only a few tens of micrometers. But in a living cell, accessory proteins (see below) crosslink and bundle them together, making these large-scale actin structures much stronger than an individual actin filament.

Microtubules and Actin Filaments Have Two Distinct Ends That Grow at Different Rates

The regular, parallel orientation of their subunits gives actin filaments and microtubules structural polarity. This orientation makes the two ends of each polymer different in ways that have a profound effect on filament growth rates. Addition of a subunit to either end of a filament of n subunits results in a filament of $n + 1$ subunits. In the absence of ATP or GTP hydrolysis, the free energy difference, and therefore the equilibrium constant (and the critical concentration), must be the same for addition of subunits at either end of the polymer. In this case, the ratio of the forward and backward rate constants, k_{on}/k_{off}, must be identical at the two ends, even though the absolute values of these rate constants may be very different at each end.

In a structurally polar filament, the kinetic rate constants for association and dissociation—k_{on} and k_{off}, respectively—are often much greater at one end than at the other. Thus, if an excess of purified subunits is allowed to assemble onto marked fragments of preformed filaments, one end of each fragment elongates much faster than the other (**Figure 16–13**). If filaments are rapidly diluted so that the free subunit concentration drops below the critical concentration, the fast-growing end also depolymerizes fastest. The more dynamic of the two ends of a filament, where both growth and shrinkage are fast, is called the **plus end**, and the other end is called the **minus end**.

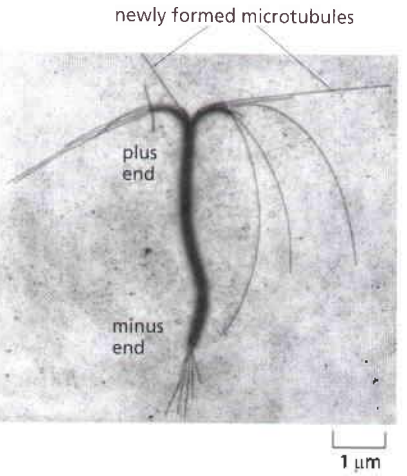


Figure 16–13 The preferential growth of microtubules at the plus end. Both microtubules and actin filaments grow faster at one end than at the other. In this case, a stable bundle of microtubules obtained from the core of a cilium (discussed later) was incubated for a short time with tubulin subunits under polymerizing conditions. Microtubules grow fastest from the plus end of the microtubule bundle, the end at the *top* in this micrograph. (Courtesy of Gary Borisy.)

On microtubules, α subunits are exposed at the minus end, and β subunits are exposed at the plus end. On actin filaments, the ATP-binding cleft on the monomer points toward the minus end. (For historical reasons, the plus ends of actin filaments are usually referred to as "barbed" ends, and minus ends as "pointed" ends, because of the arrowhead appearance of myosin heads when bound along the filament.)

Filament elongation proceeds spontaneously when the free energy change (ΔG) for addition of the soluble subunit is less than zero. This is the case when the concentration of subunits in solution exceeds the critical concentration. Likewise, filament depolymerization proceeds spontaneously when this free energy change is greater than zero. A cell can couple an energetically unfavorable process to these spontaneous processes; thus, the cell can use free energy released during spontaneous filament polymerization or depolymerization to do mechanical work—in particular, to push or pull an attached load. For example, elongating microtubules can help push out membranes, and shrinking microtubules can help pull mitotic chromosomes away from their sisters during anaphase. Similarly, elongating actin filaments help protrude the leading edge of motile cells, as we discuss later.

Filament Treadmilling and Dynamic Instability Are Consequences of Nucleotide Hydrolysis by Tubulin and Actin

Thus far, our discussion of filament dynamics has ignored a critical fact that applies to both actin filaments and microtubules. In addition to their ability to form noncovalent polymers, the actin and tubulin subunits are both enzymes that can catalyze the hydrolysis of a nucleoside triphosphate, ATP or GTP, respectively. For the free subunits, this hydrolysis proceeds very slowly; however, it is accelerated when the subunits are incorporated into filaments. Shortly after incorporation of an actin or tubulin subunit into a filament, nucleotide hydrolysis occurs; the free phosphate group is released from each subunit, but the nucleoside diphosphate remains trapped in the filament structure. (On tubulin, the nucleotide-binding site lies at the interface between two neighboring subunits—see Figure 16–11, whereas in actin, the nucleotide is deep in a cleft near the center of the subunit—see Figure 16–12.) Thus, two different types of filament structures can exist, one with the "T form" of the nucleotide bound (ATP for actin, GTP for tubulin), and one with the "D form" bound (ADP for actin, GDP for tubulin).

When the nucleotide is hydrolyzed, much of the free energy released by cleavage of the high-energy phosphate–phosphate bond is stored in the polymer lattice. This makes the free energy change for dissociation of a subunit from the D-form polymer more negative than the free energy change for dissociation of a subunit from the T-form polymer. Consequently, the ratio of k_{off}/k_{on} for the D-form polymer, which is numerically equal to its critical concentration [C_c(D)], is larger than the corresponding ratio for the T-form polymer. Thus, C_c(D) is greater than C_c(T). For certain concentrations of free subunits, D-form polymers will therefore shrink while T-form polymers grow.

In living cells, most of the free subunits are in the T form, as the free concentration of both ATP and GTP is about ten-fold higher than that of ADP and GDP. The longer the time that subunits have been in the polymer lattice, the more likely they are to have hydrolyzed their bound nucleotide. Whether the subunit at the very end of a filament is in the T or the D form depends on the rate of this hydrolysis compared with the rate of subunit addition. If the rate of subunit addition is high, that is if the filament is growing rapidly, then it is likely that

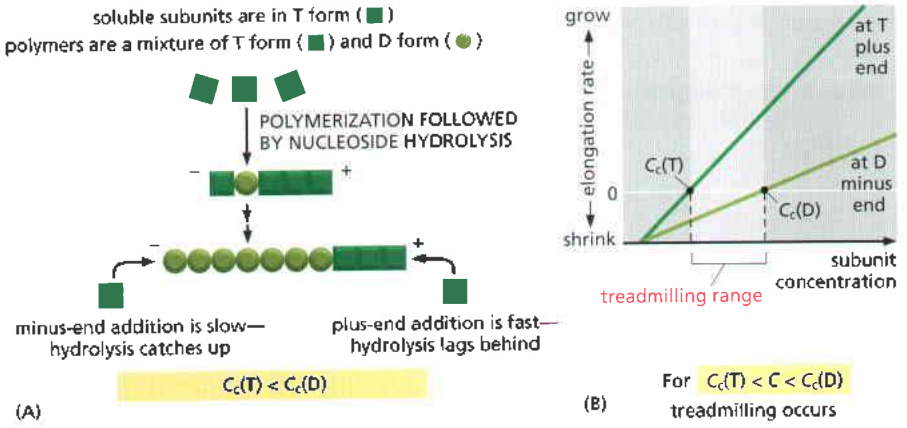


Figure 16–14 The treadmilling of an actin filament or microtubule, made possible by the nucleoside triphosphate hydrolysis that follows subunit addition. (A) Explanation for the different critical concentrations (C_c) at the plus and minus ends. Subunits with bound nucleoside triphosphate (T-form subunits) polymerize at both ends of a growing filament, and then undergo nucleotide hydrolysis in the filament lattice. As the filament grows, elongation is faster than hydrolysis at the plus end in this example, and the terminal subunits at this end are therefore always in the T form. However, hydrolysis is faster than elongation at the minus end, and so terminal subunits at this end are in the D form. (B) Treadmilling occurs at intermediate concentrations of free subunits. The critical concentration for polymerization on a filament end in the T form is lower than for a filament end in the D form. If the actual subunit concentration is somewhere between these two values, the plus end grows while the minus end shrinks, resulting in treadmilling.

a new subunit will add on to the polymer before the nucleotide in the previously added subunit has been hydrolyzed, so that the tip of the polymer remains in the T form, forming an *ATP cap* or *GTP cap*. However, if the rate of subunit addition is low, hydrolysis may occur before the next subunit is added, and the tip of the filament will then be in the D form.

The rate of subunit addition at the end of a filament is the product of the free subunit concentration and the rate constant k_{on}. The k_{on} is much faster for the plus end of a filament than for the minus end because of a structural difference between the two ends (see Panel 16–2). At an intermediate concentration of free subunits, it is therefore possible for the rate of subunit addition to be faster than nucleotide hydrolysis at the plus end, but slower than nucleotide hydrolysis at the minus end. In this case, the plus end of the filament remains in the T conformation, while the minus end adopts the D conformation. As just explained, the D form has a higher critical concentration than the T form. (In other words, the D form leans more readily toward disassembly, while the T form leans more readily toward assembly). If the concentration of free subunits in solution is in an intermediate range—higher than the critical concentration of the T form (that is, the plus end), but lower than the critical concentration of the D form (that is, the minus end)—the filament adds subunits at the plus end, and simultaneously loses subunits from the minus end. This leads to the remarkable property of filament **treadmilling** (**Figure 16–14** and Panel 16–2).

During treadmilling, subunits are recruited at the plus end of the polymer in the T form and shed from the minus end in the D form. The ATP or GTP hydrolysis that occurs along the way gives rise to the difference in the free energy of the association/dissociation reactions at the plus and minus ends of the actin filament or microtubule and thereby makes treadmilling possible. At a particular intermediate subunit concentration, the filament growth at the plus end exactly balances the filament shrinkage at the minus end. Now, the subunits cycle rapidly between the free and filamentous states, while the total length of the filament remains unchanged. This "steady-state treadmilling" requires a constant consumption of energy in the form of nucleoside triphosphate hydrolysis. While the extent of treadmilling inside the cell is uncertain, the treadmilling of single filaments has been observed *in vitro* for actin, and a phenomenon that looks like treadmilling can be observed in live cells for individual microtubules (**Figure 16–15**).

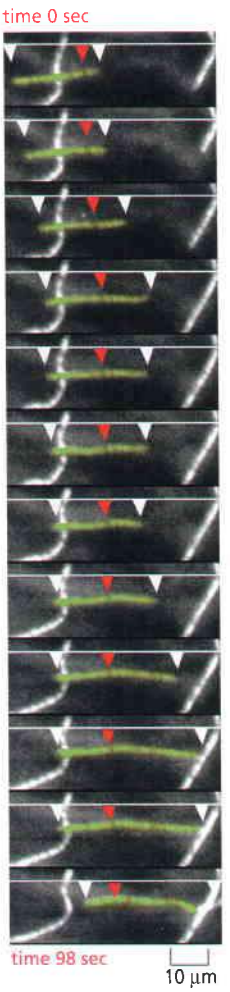
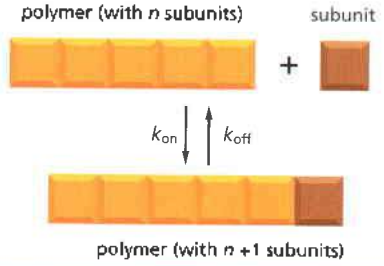


Figure 16–15 Treadmilling behavior of a microtubule, as observed in a living cell. A cell was injected with tubulin that had been covalently linked to the fluorescent dye rhodamine, so that approximately 1 tubulin subunit in 20 was fluorescent. The fluorescence of individual microtubules was then observed with a sensitive electronic camera. The microtubule shown appears to be sliding from *left* to *right*, but, in fact, the microtubule lattice remains stationary (as shown by a landmark within the microtubule lattice indicated by the *red* arrowhead), while the plus end (on the *right*) grows and the minus end (on the *left*) shrinks. The plus end also displays dynamic instability. (From C.M. Waterman-Storer and E.D. Salmon, *J. Cell Biol.* 139:417–434, 1997. With permission from The Rockefeller University Press.)

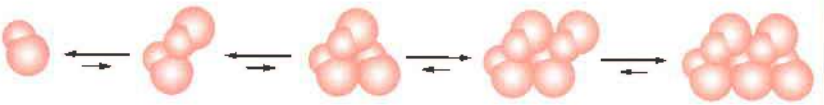
ON RATES AND OFF RATES

A linear polymer of protein molecules, such as an actin filament or a microtubule, assembles (polymerizes) and disassembles (depolymerizes) by the addition and removal of subunits at the ends of the polymer. The rate of addition of these subunits (called monomers) is given by the rate constant k_{on}, which has units of M^{-1} sec^{-1}. The rate of loss is given by k_{off} (units of sec^{-1}).



NUCLEATION

A helical polymer is stabilized by multiple contacts between adjacent subunits. In the case of actin, two actin molecules bind relatively weakly to each other, but addition of a third actin monomer to form a trimer makes the entire group more stable.



Further monomer addition can take place onto this trimer, which therefore acts as a nucleus for polymerization. For tubulin, the nucleus is larger and has a more complicated structure (possibly a ring of 13 or more tubulin molecules)—but the principle is the same.

The assembly of a nucleus is relatively slow, which explains the lag phase seen during polymerization. The lag phase can be reduced or abolished entirely by adding premade nuclei, such as fragments of already polymerized microtubules or actin filaments.

THE CRITICAL CONCENTRATION

The number of monomers that add to the polymer (actin filament or microtubule) per second will be proportional to the concentration of the free subunit ($k_{on}C$), but the subunits will leave the polymer end at a constant rate (k_{off}) that does not depend on C. As the polymer grows, subunits are used up, and C is observed to drop until it reaches a constant value, called the critical concentration (C_c). At this concentration the rate of subunit addition equals the rate of subunit loss.

At this equilibrium,

$$k_{on}\, C = k_{off}$$

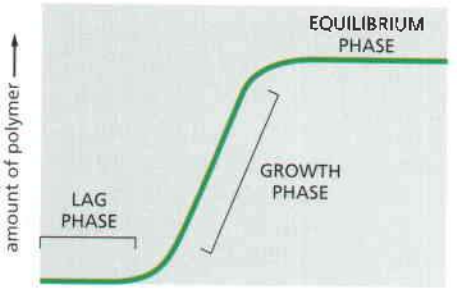
so that

$$C_c = \frac{k_{off}}{k_{on}} = \frac{1}{K}$$

(where K is the equilibrium constant for subunit addition; see Figure 3–43).

TIME COURSE OF POLYMERIZATION

The assembly of a protein into a long helical polymer such as a cytoskeletal filament or a bacterial flagellum typically shows the following time course:



The lag phase corresponds to time taken for nucleation.

The growth phase occurs as monomers add to the exposed ends of the growing filament, causing filament elongation.

The equilibrium phase, or steady state, is reached when the growth of the polymer due to monomer addition precisely balances the shrinkage of the polymer due to disassembly back to monomers.

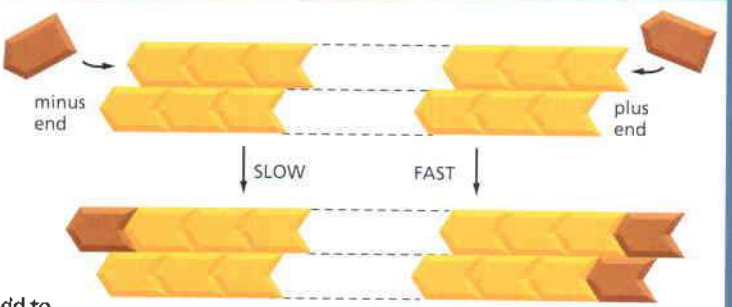
PLUS AND MINUS ENDS

The two ends of an actin filament or microtubule polymerize at different rates. The fast-growing end is called the plus end, whereas the slow-growing end is called the minus end. The difference in the rates of growth at the two ends is made possible by changes in the conformation of each subunit as it enters the polymer.



This conformational change affects the rates at which subunits add to the two ends.

Even though k_{on} and k_{off} will have different values for the plus and minus ends of the polymer, their ratio k_{off}/k_{on}—and hence C_c—must be the same at both ends for a simple polymerization reaction (no ATP or GTP hydrolysis). This is because exactly the same subunit interactions are broken when a subunit is lost at either end, and the final state of the subunit after dissociation is identical. Therefore, the ΔG for subunit loss, which determines the equilibrium constant for its association with the end, is identical at both ends: if the plus end grows four times faster than the minus end, it must also shrink four times faster. Thus, for $C > C_c$, both ends grow; for $C < C_c$, both ends shrink.



The nucleoside triphosphate hydrolysis that accompanies actin and tubulin polymerization removes this constraint.

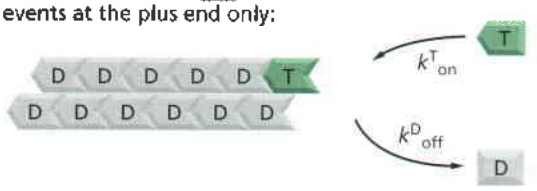
NUCLEOTIDE HYDROLYSIS

Each actin molecule carries a tightly bound ATP molecule that is hydrolyzed to a tightly bound ADP molecule soon after its assembly into the polymer. Similarly, each tubulin molecule carries a tightly bound GTP that is converted to a tightly bound GDP molecule soon after the molecule assembles into the polymer.



Hydrolysis of the bound nucleotide reduces the binding affinity of the subunit for neighboring subunits and makes it more likely to dissociate from each end of the filament (see Figure 16–16 for a possible mechanism). It is usually the T form that adds to the filament and the D form that leaves.

Considering events at the plus end only:



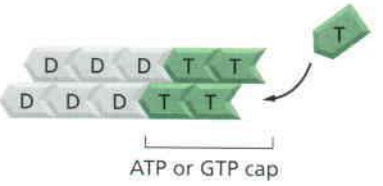
As before, the polymer will grow until $C = C_c$. For illustrative purposes, we can ignore k^D_{on} and k^T_{off} since they are usually very small, so that polymer growth ceases when

$$k^T_{on}\, C = k^D_{off} \quad \text{or} \quad C_c = \frac{k^D_{off}}{k^T_{on}}$$

This is a steady state and not a true equilibrium, because the ATP or GTP that is hydrolyzed must be replenished by a nucleotide exchange reaction of the free subunit (D → T).

ATP CAPS AND GTP CAPS

The rate of addition of subunits to a growing actin filament or microtubule can be faster than the rate at which their bound nucleotide is hydrolyzed. Under such conditions, the end has a "cap" of subunits containing the nucleoside triphosphate—an ATP cap on an actin filament or a GTP cap on a microtubule.



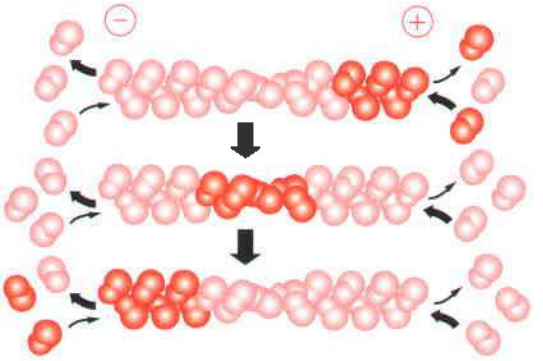
DYNAMIC INSTABILITY and TREADMILLING are two behaviors observed in cytoskeletal polymers. Both are associated with nucleoside triphosphate hydrolysis. Dynamic instability is believed to predominate in microtubules, whereas treadmilling may predominate in actin filaments.

TREADMILLING

One consequence of the nucleotide hydrolysis that accompanies polymer formation is to change the critical concentration at the two ends of the polymer. Since k^D_{off} and k^T_{on} refer to different reactions, their ratio k^D_{off}/k^T_{on} need not be the same at both ends of the polymer, so that:

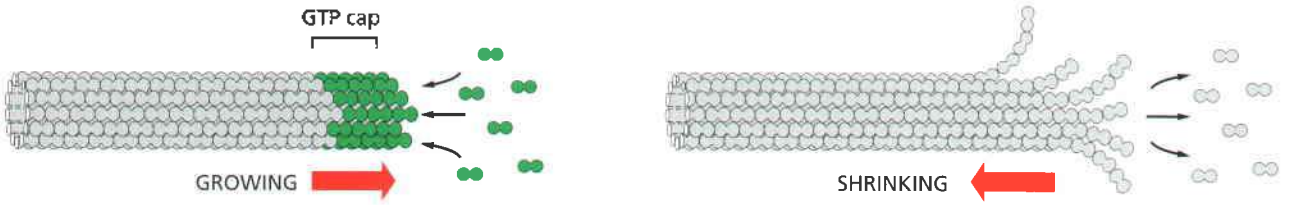
$$C_c \text{ (minus end)} > C_c \text{ (plus end)}$$

Thus, if both ends of a polymer are exposed, polymerization proceeds until the concentration of free monomer reaches a value that is above C_c for the plus end but below C_c for the minus end. At this steady state, subunits undergo a net assembly at the plus end and a net disassembly at the minus end at an identical rate. The polymer maintains a constant length, even though there is a net flux of subunits through the polymer, known as treadmilling.



DYNAMIC INSTABILITY

Microtubules depolymerize about 100 times faster from an end containing GDP tubulin than from one containing GTP tubulin. A GTP cap favors growth, but if it is lost, then depolymerization ensues.



Individual microtubules can therefore alternate between a period of slow growth and a period of rapid disassembly, a phenomenon called dynamic instability.

The kinetic differences between the behavior of the T form and the D form have another important consequence for the behaviors of filaments. If the rate of subunit addition at one end is similar in magnitude to the rate of hydrolysis, there is a finite probability that this end will start out in a T form, but that hydrolysis will eventually "catch up" with the addition and transform the end to a D form. This transformation is sudden and random, with a certain probability per unit time.

Suppose that the concentration of free subunits is intermediate between the critical concentration for a T-form end and the critical concentration for a D-form end (that is, in the same range of concentrations where treadmilling is observed). Now, any end that happens to be in the T form will grow, whereas any end that happens to be in the D form will shrink. On a single filament, an end might grow for a certain length of time in a T form, but then suddenly change to the D form and begin to shrink rapidly, even while the free subunit concentration is held constant. At some later time, it might then regain a T-form end and begin to grow again. This rapid interconversion between a growing and shrinking state, at a uniform free subunit concentration, is called **dynamic instability** (**Figure 16–16A**). The change from growth to rapid shrinkage is called a *catastrophe*, while the change to growth is called a *rescue*. <CCCA>

In a population of microtubules, at any instant some of the ends are in the T form and some are in the D form, with the ratio depending on the hydrolysis rate and the free subunit concentration. The structural difference between a T-form end and a D-form end is dramatic. Tubulin subunits with GTP bound to the β-monomer produce straight protofilaments that make strong and regular lateral contacts with one another. But the hydrolysis of GTP to GDP is associated with a subtle conformational change in the protein, which makes the protofilaments curved (Figure 16–16B). On a rapidly growing microtubule, the GTP cap is thought to constrain the curvature of the protofilaments, and the ends appear straight. But when the terminal subunits have hydrolyzed their nucleotides, this constraint is removed, and the curved protofilaments spring apart. This cooperative release of the energy of hydrolysis stored in the microtubule lattice causes the curled protofilaments to peel off rapidly, and rings and curved oligomers of GDP-containing tubulin are seen near the ends of depolymerizing microtubules (Figure 16–16C).

Actin filaments also undergo length fluctuations but on a much smaller scale, so that at steady state the length fluctuates only a micrometer or so over several minutes, as compared to tens of micrometers for microtubules undergoing dynamic instability. In most eucaryotic cells, dynamic instability is thought to predominate in microtubules, whereas treadmilling may predominate in actin filaments.

Treadmilling and Dynamic Instability Aid Rapid Cytoskeletal Rearrangement <AAAT>

Both dynamic instability and treadmilling allow a cell to maintain the same overall filament content, while individual subunits constantly recycle between the filaments and the cytosol. How dynamic are the microtubules and actin filaments inside a living cell? Typically, a microtubule, with major structural differences between its growing and shrinking ends, switches between growth and shrinkage every few minutes. The ends of individual microtubules can therefore be seen in real time to exhibit dynamic instability (**Figure 16–17**). Because of their smaller size and denser packing, it is much more difficult to resolve the ends of individual actin filaments within living cells. With appropriate techniques based on fluorescence microscopy, however, one can show that actin filament turnover is typically rapid, with individual filaments persisting for only a few tens of seconds or minutes.

At first glance, the dynamic behavior of filaments seems like a waste of energy. To maintain a constant concentration of actin filaments and microtubules, most of which are undergoing a process of either treadmilling or dynamic instability, the cell must hydrolyze large amounts of nucleoside

triphosphate. As we explained with our ant-trail analogy at the beginning of the chapter, the advantage to the cell seems to be the spatial and temporal flexibility that is inherent in a structural system with constant turnover. Individual subunits are small and can diffuse very rapidly; an actin or tubulin subunit can diffuse across the diameter of a typical eucaryotic cell in several seconds. As noted

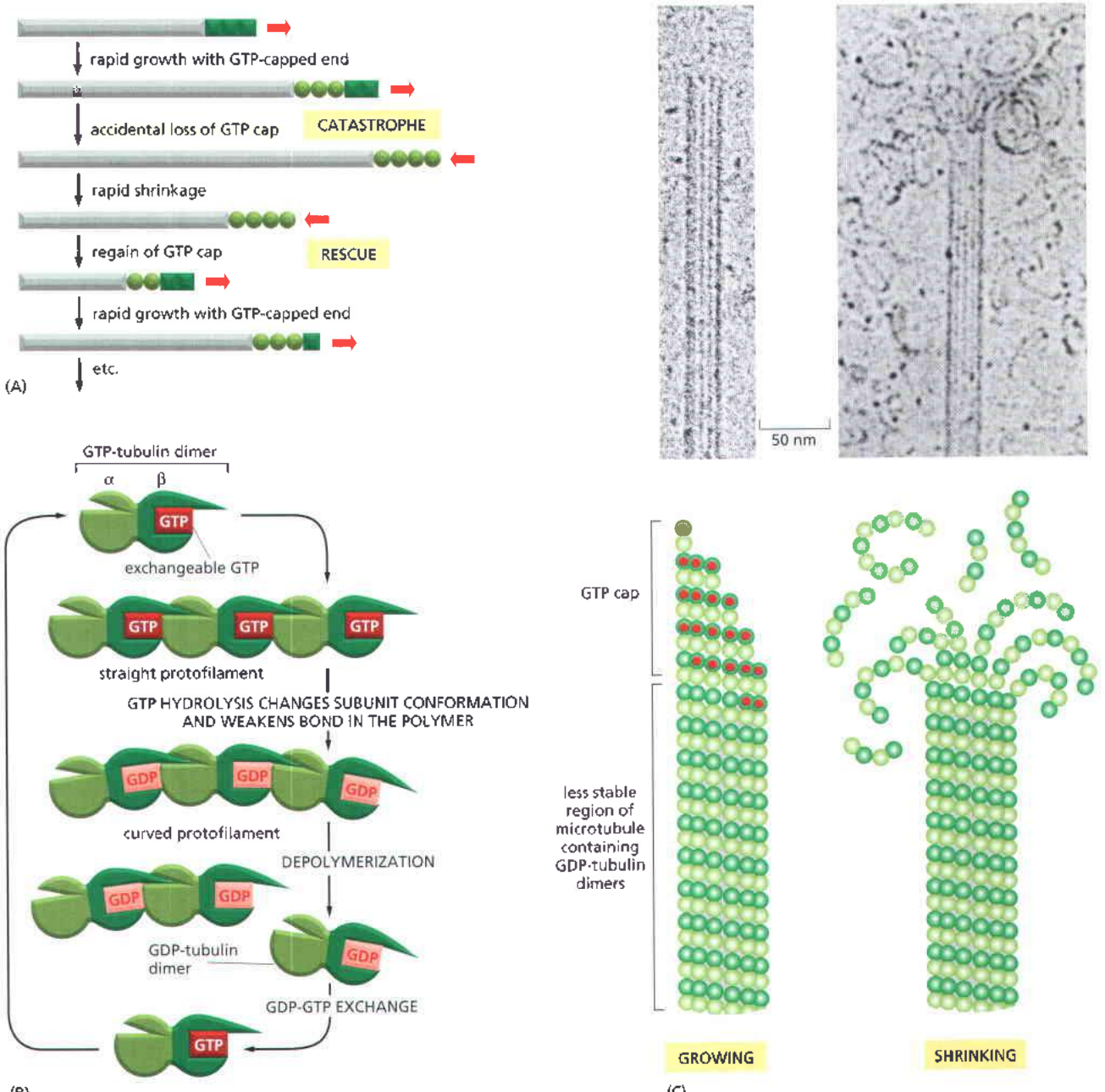


Figure 16–16 Dynamic instability due to the structural differences between a growing and a shrinking microtubule end. (A) If the free tubulin concentration in solution is between the critical values indicated in Figure 16–14B, a single microtubule end may undergo transitions between a growing state and a shrinking state. A growing microtubule has GTP-containing subunits at its end, forming a GTP cap. If nucleotide hydrolysis proceeds more rapidly than subunit addition, the cap is lost and the microtubule begins to shrink, an event called a "catastrophe." But GTP-containing subunits may still add to the shrinking end, and if enough add to form a new cap, then microtubule growth resumes, an event called "rescue." (B) Model for the structural consequences of GTP hydrolysis in the microtubule lattice. The addition of GTP-containing tubulin subunits to the end of a protofilament causes the end to grow in a linear conformation that can readily pack into the cylindrical wall of the microtubule. Hydrolysis of GTP after assembly changes the conformation of the subunits and tends to force the protofilament into a curved shape that is less able to pack into the microtubule wall. (C) In an intact microtubule, protofilaments made from GDP-containing subunits are forced into a linear conformation by the many lateral bonds within the microtubule wall, given a stable cap of GTP-containing subunits. Loss of the GTP cap, however, allows the GDP-containing protofilaments to relax into their more curved conformation. This leads to a progressive disruption of the microtubule. Above the drawings of a growing and a shrinking microtubule, electron micrographs show actual microtubules in each of these two states, as observed in preparations in vitreous ice. Note particularly the curling, disintegrating GDP-containing protofilaments at the end of the shrinking microtubule. (C, courtesy of E.M. Mandelkow, E. Mandelkow and R.A. Milligan, *J. Cell Biol.* 114:977–991, 1991. With permission from The Rockefeller University Press.)

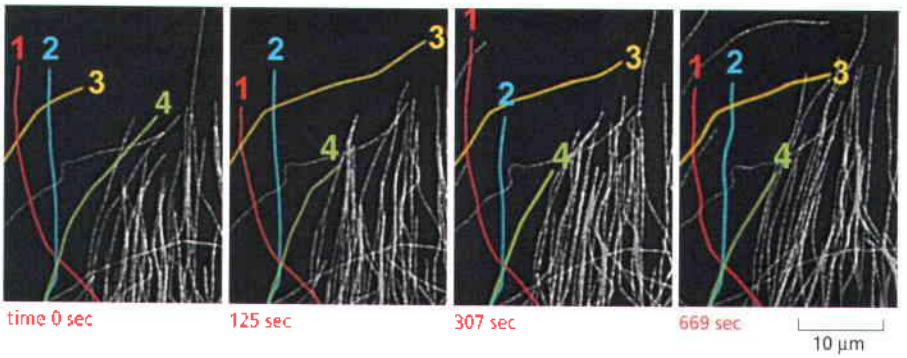


Figure 16–17 Direct observation of the dynamic instability of microtubules in a living cell. <AAAT> Microtubules in a newt lung epithelial cell were observed after the cell was injected with a small amount of rhodamine labeled tubulin, as in Figure 16–15. Notice the dynamic instability of microtubules at the edge of the cell. Four individual microtubules are highlighted for clarity; each of these shows alternating shrinkage and growth. (Courtesy of Wendy C. Salmon and Clare Waterman-Storer.)

previously, the rate-limiting step in the formation of a new filament is nucleation, so these rapidly diffusing subunits tend to assemble either on the ends of preexisting filaments or at particular sites where special proteins catalyze the nucleation step. The new filaments in either case are highly dynamic, and unless specifically stabilized, they have only a fleeting existence. By controlling where filaments are nucleated and selectively stabilized, a cell can control the location of its filament systems, and hence its structure. It seems that the cell is continually testing a wide variety of internal structures and only preserving those that are useful. When external conditions change, or when internal signals arise (as during the transitions in the cell cycle), the cell is poised to change its structure rapidly (see Figures 16–2 to 16–4).

Actin and tubulin have independently evolved their nucleoside triphosphate hydrolysis to enable their filaments to depolymerize readily after they have polymerized. These two proteins are completely unrelated in amino acid sequence: actin is distantly related in structure to the glycolytic enzyme hexokinase, whereas tubulin is distantly related to a large family of GTPases that includes the heterotrimeric G proteins and monomeric GTPases such as Ras (discussed in Chapter 3). In both protein families, the coupling between nucleotide hydrolysis and a protein conformational change that alters protein function appears to be evolutionarily very ancient; however, the purposes of that structural coupling have diverged over time to include signal transmission, catalysis, and regulation of the polymerization/depolymerization cycle.

In certain specialized structures, parts of the cytoskeleton become less dynamic. In a terminally differentiated cell such as a neuron, for example, it is desirable to maintain a consistent structure over time, and many of the actin filaments and microtubules are stabilized by association with other proteins. However, when new connections are made in the brain, as when the information you are reading now is transferred into long-term memory, even a cell as stable as a neuron can grow new elongated processes to make new synapses. To do this, a neuron requires the dynamic, exploratory activities of its cytoskeletal filaments.

Tubulin and Actin Have Been Highly Conserved During Eucaryotic Evolution

Tubulin is found in all eucaryotic cells, and it exists in multiple isoforms. Yeast and human tubulins are 75% identical in amino acid sequence. In mammals, there are at least six forms of α-tubulin and a similar number of forms of β-tubulin, each encoded by a different gene. The different forms of tubulin are very similar, and they generally will copolymerize into mixed microtubules in the test tube. However, they can have distinct locations in a cell and perform subtly different functions. As a striking example, a specific form of β-tubulin forms the microtubules in six specialized touch-sensitive neurons in the nematode *Caenorhabditis elegans*. Mutations that eliminate this protein result in the loss of touch-sensitivity, with no apparent defect in other functions.

Like tubulin, actin is found in all eucaryotic cells. Most organisms have multiple genes encoding actin; humans have six. Actin is extraordinarily well conserved among eucaryotes. The amino acid sequences of actins from different species are usually about 90% identical. But, again like tubulin, small variations in actin amino acid sequence can cause significant functional differences. In vertebrates, there are three subtly different isoforms of actin, termed α, β, and γ, that differ slightly in their amino acid sequences. The α-actin is expressed only in muscle cells, while β and γ are found together in almost all nonmuscle cells. Yeast actin and *Drosophila* muscle actin are 89% identical, yet the expression of yeast actin in *Drosophila* results in a fly that looks normal but is unable to fly.

Why are the amino acid sequences of actin and tubulin so strictly conserved in eucaryotic evolution, whereas the sequences of most other cytoskeletal proteins, including intermediate filament proteins and the large families of accessory proteins that bind to actin or tubulin, are not? The likely explanation is that the need for large numbers of other proteins to interact with the entire surface of an actin filament or microtubule limits the variability of their structures. Genetic and biochemical studies in the yeast *Saccharomyces cerevisiae* have demonstrated that actin interacts directly with dozens of other proteins, and indirectly with even more (**Figure 16–18**). Thus, any mutation in actin that could result in a desirable change in its interaction with one protein might cause undesirable changes in its interactions with other proteins that bind at or near the same site. Over time, evolving organisms have found it more profitable to leave actin and tubulin alone, and to alter their binding partners instead.

Intermediate Filament Structure Depends on the Lateral Bundling and Twisting of Coiled Coils

All eucaryotic cells contain actin and tubulin. But the third major type of cytoskeletal protein, the *intermediate filament*, forms a cytplasmic filament in only some metazoans—including vertebrates, nematodes, and mollusks. Even in these organisms, intermediate filaments are not required in the cytoplasm of every cell type. The specialized glial cells (called oligodendrocytes) that make

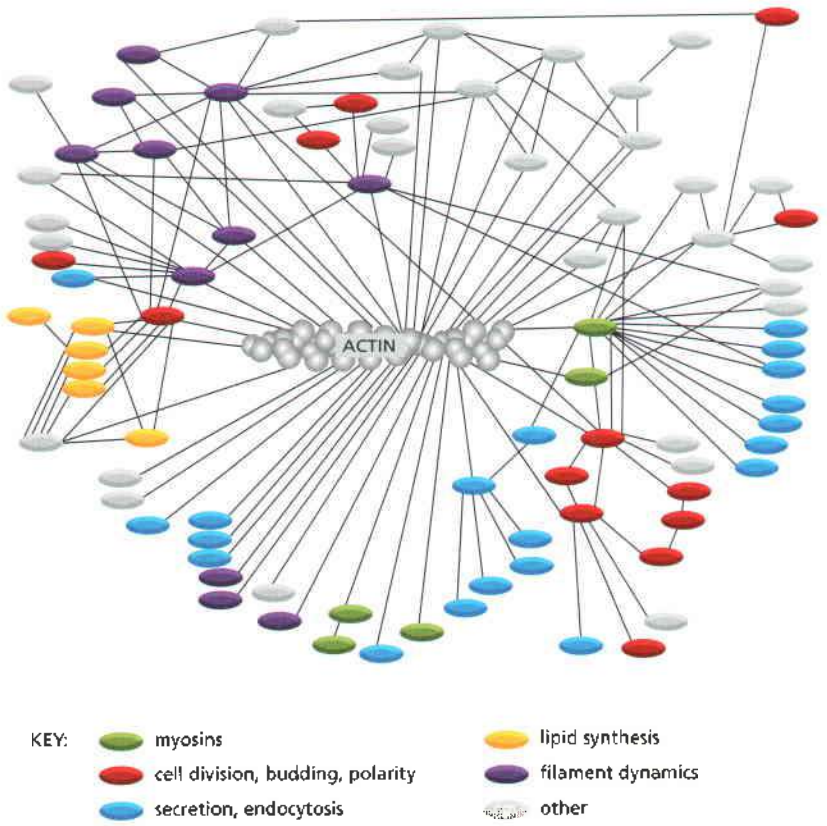


Figure 16–18 Actin at the crossroads. Actin binds to a very large variety of accessory proteins in all eucaryotic cells. This diagram shows most of the interactions that have been demonstrated, using either genetic or biochemical techniques, in the yeast *Saccharomyces cerevisiae*. Accessory proteins that operate in the same intracellular process are shown in the same color, as indicated in the key. (Adapted from D. Botstein et al., in The Molecular and Cellular Biology of the Yeast *Saccharomyces* [J.R. Broach, J.R. Pringle, E.W. Jones, eds.], Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1991.)

myelin in the vertebrate central nervous system, for example, do not contain such intermediate filaments. Intermediate filaments are particularly prominent in the cytoplasm of cells that are subject to mechanical stress, and are generally not found in animals that have rigid exoskeletons such as arthropods and echinoderms. It seems that intermediate filaments play an important role in imparting mechanical strength to tissues for the squishier animals.

Cytoplasmic intermediate filaments are closely related to their ancestors, the much more prevalent *nuclear lamins*. The nuclear lamins are intermediate filament proteins that form a meshwork lining the inner membrane of the eucaryotic nuclear envelope, where they provide anchorage sites for chromosomes and nuclear pores (their dynamic behavior during cell division is discussed in Chapter 12). Several times during metazoan evolution, lamin genes have apparently duplicated, and the duplicates have evolved to produce rope-like, cytoplasmic intermediate filaments.

The individual polypeptides of **intermediate filaments** are elongated molecules with an extended central α-helical domain that forms a parallel coiled coil with another monomer. A pair of parallel dimers then associates in an antiparallel fashion to form a staggered tetramer. This tetramer represents the soluble subunit that is analogous to the αβ-tubulin dimer, or to the actin monomer (**Figure 16–19**). Unlike the actin or tubulin, the intermediate filament subunits do not contain a binding site for a nucleoside triphosphate.

Figure 16–19 A model of intermediate filament construction. <GCCA> The monomer shown in (A) pairs with an identical monomer to form a dimer (B), in which the conserved central rod domains are aligned in parallel and wound together into a coiled coil. (C) Two dimers then line up side by side to form an antiparallel tetramer of four polypeptide chains. The tetramer is the soluble subunit of intermediate filaments. (D) Within each tetramer, the two dimers are offset with respect to one another, thereby allowing it to associate with another tetramer. (E) In the final 10-nm rope-like filament, tetramers are packed together in a helical array, which has 16 dimers (32 coiled coils) in cross-section. Half of these dimers are pointing in each direction. An electron micrograph of intermediate filaments is shown on the upper left. (Electron micrograph courtesy of Roy Quinlan.)

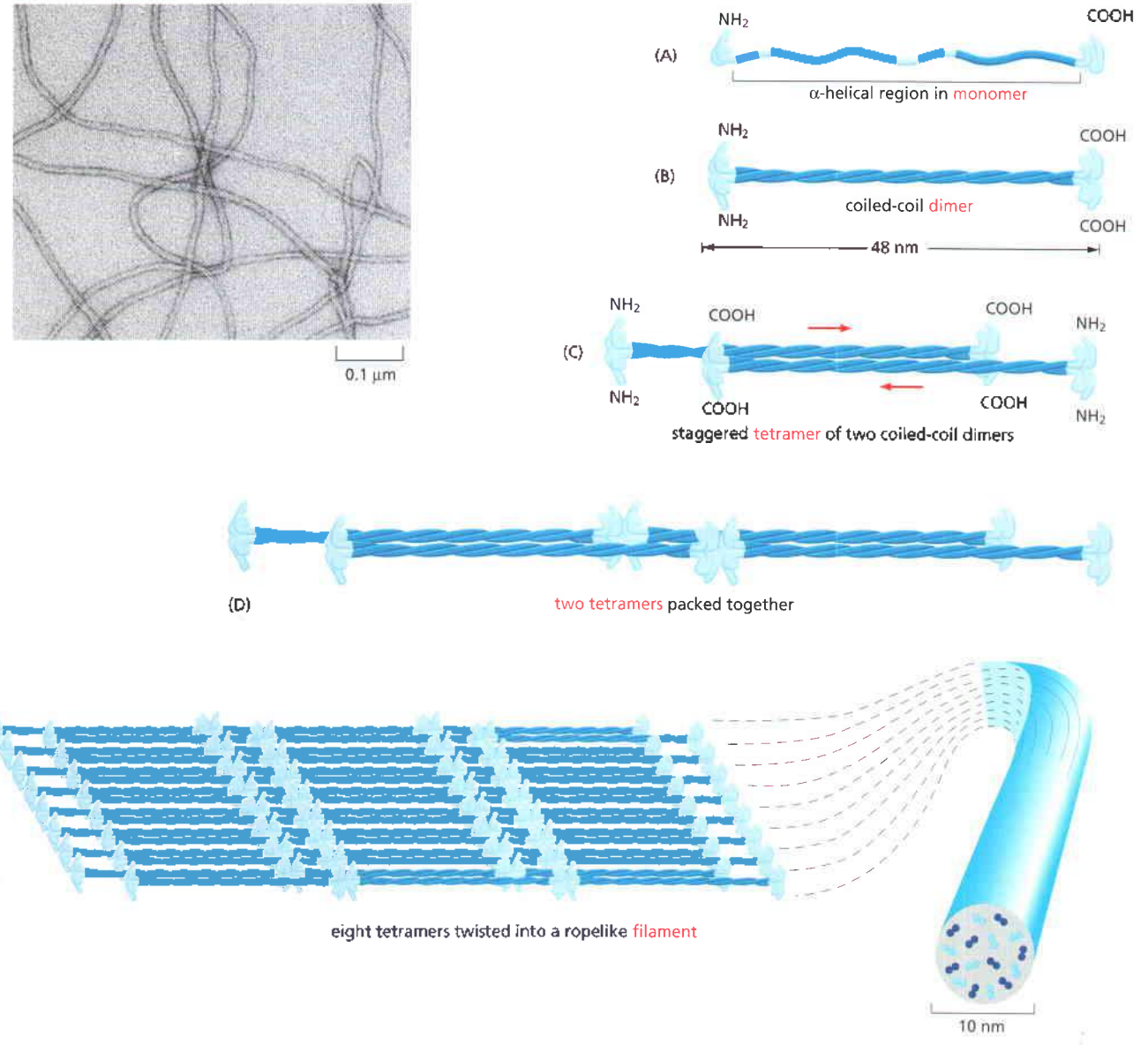


Table 16–1 Major Types of Intermediate Filament Proteins in Vertebrate Cells

TYPES OF IF	COMPONENT POLYPEPTIDES	LOCATION
Nuclear	lamins A, B, and C	nuclear lamina (inner lining of nuclear envelope)
Vimentin-like	vimentin	many cells of mesenchymal origin
	desmin	muscle
	glial fibrillary acidic protein	glial cells (astrocytes and some Schwann cells)
	peripherin	some neurons
Epithelial	type I keratins (acidic)	epithelial cells and their derivatives (e.g., hair and nails)
	type II keratins (basic)	
Axonal	neurofilament proteins (NF-L, NF-M, and NF-H)	neurons

Since the tetrameric subunit is made up of two dimers pointing in opposite directions, its two ends are the same. The assembled intermediate filament therefore lacks the overall structural polarity that is critical for actin filaments and microtubules. The tetramers pack together laterally to form the filament, which includes eight parallel protofilaments made up of tetramers. Each individual intermediate filament therefore has a cross section of 32 individual α-helical coils. This large number of polypeptides all lined up together, with the strong lateral hydrophobic interactions typical of coiled-coil proteins, gives intermediate filaments a rope-like character. They can be easily bent, with a persistence length of less than one micrometer (compared to several millimeters for microtubules and about ten micrometers for actin), but they are extremely difficult to break.

Less is understood about the mechanism of assembly and disassembly of intermediate filaments than of actin filaments and microtubules, but some types of intermediate filaments including *vimentin* form highly dynamic structures in cells such as fibroblasts. Under normal conditions, protein phosphorylation probably regulates their disassembly, in much the same way that phosphorylation regulates the disassembly of nuclear lamins in mitosis (see Figure 12–20). As evidence for rapid turnover, labeled subunits microinjected into tissue culture cells rapidly add themselves onto the existing intermediate filaments within a few minutes, while an injection of peptides derived from a conserved helical region of the subunit induces the rapid disassembly of the intermediate filament network. Interestingly, the latter injection can also induce the disassembly of the microtubule and actin filament networks in some cases, demonstrating that there is a fundamental mechanical integration of the three cytoskeletal systems in these cells.

Intermediate Filaments Impart Mechanical Stability to Animal Cells

Intermediate filaments come in a wide variety of types, with substantially more sequence variation in the subunit isoforms than occurs in the isoforms of actin or tubulin. A central α-helical domain has 40 or so heptad repeat motifs that form an extended coiled-coil structure (see Figure 3–9). This domain is similar in the different isoforms, but the N- and C-terminal globular domains can vary greatly.

Different families of intermediate filaments are expressed in different cell types (**Table 16–1**). **Keratins** are the most diverse intermediate filament family: there are about 20 found in different types of human epithelial cells, and about 10 more that are specific to hair and nails; analysis of the human genome sequence has revealed that there may be about 50 distinct keratins. Every keratin filament is made up of an equal mixture of type I (acidic) and type II (neutral/basic) keratin chains; these form heterodimers, two of which then join

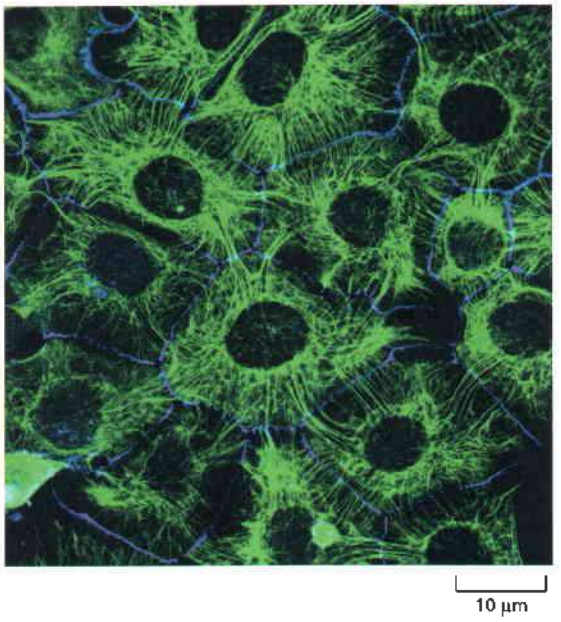


Figure 16–20 Keratin filaments in epithelial cells. Immunofluorescence micrograph of the network of keratin filaments *(green)* in a sheet of epithelial cells in culture. The filaments in each cell are indirectly connected to those of its neighbors by desmosomes (discussed in Chapter 19). A second protein *(blue)* has been stained to reveal the location of the cell boundaries. (Courtesy of Kathleen Green and Evangeline Amargo.)

to form the fundamental tetrameric subunit (see Figure 16–19). Cross-linked keratin networks held together by disulfide bonds may survive even the death of their cells, forming tough coverings for animals, as in the outer layer of skin and in hair, nails, claws, and scales. The diversity in keratins is clinically useful in the diagnosis of epithelial cancers (carcinomas), as the particular set of keratins expressed gives an indication of the epithelial tissue in which the cancer originated and thus can help to guide the choice of treatment.

A single epithelial cell may produce multiple types of keratins, and these copolymerize into a single network (**Figure 16–20**). Keratin filaments impart mechanical strength to epithelial tissues in part by anchoring the intermediate filaments at sites of cell-cell contact, called *desmosomes*, or cell-matrix contact, called *hemidesmosomes* (see Figure 16–5). We discuss these important adhesive structures in more detail in Chapter 19.

Mutations in keratin genes cause several human genetic diseases. For example, when defective keratins are expressed in the basal cell layer of the epidermis, they produce a disorder called *epidermolysis bullosa simplex*, in which the skin blisters in response to even very slight mechanical stress, which ruptures the basal cells (**Figure 16–21**). Other types of blistering diseases, including disorders of the mouth, esophageal lining, and the cornea of the eye, are caused by mutations in the different keratins whose expression is specific to those tissues. All of these maladies are typified by cell rupture as a consequence of mechanical trauma and a disorganization or clumping of the keratin filament cytoskeleton. Many of the specific mutations that cause these diseases alter the ends of

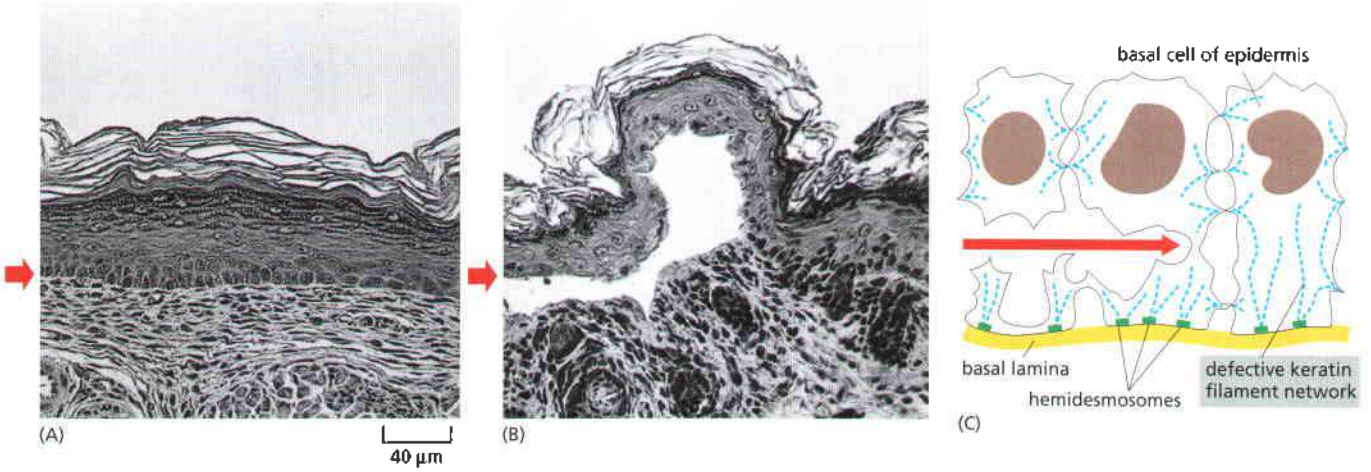


Figure 16–21 Blistering of the skin caused by a mutant keratin gene. A mutant gene encoding a truncated keratin protein (lacking both the N- and C-terminal domains) was expressed in a transgenic mouse. The defective protein assembles with the normal keratins and thereby disrupts the keratin filament network in the basal cells of the skin. Light micrographs of cross sections of normal (A) and mutant (B) skin show that the blistering results from the rupturing of cells in the basal layer of the mutant epidermis (short *red* arrows). (C) A sketch of three cells in the basal layer of the mutant epidermis, as observed by electron microscopy. As indicated by the *red* arrow, the cells rupture between the nucleus and the hemidesmosomes (discussed in Chapter 19), which connect the keratin filaments to the underlying basal lamina. (From P.A. Coulombe et al., *J. Cell Biol.* 115:1661–1674, 1991. With permission from The Rockefeller University Press.)

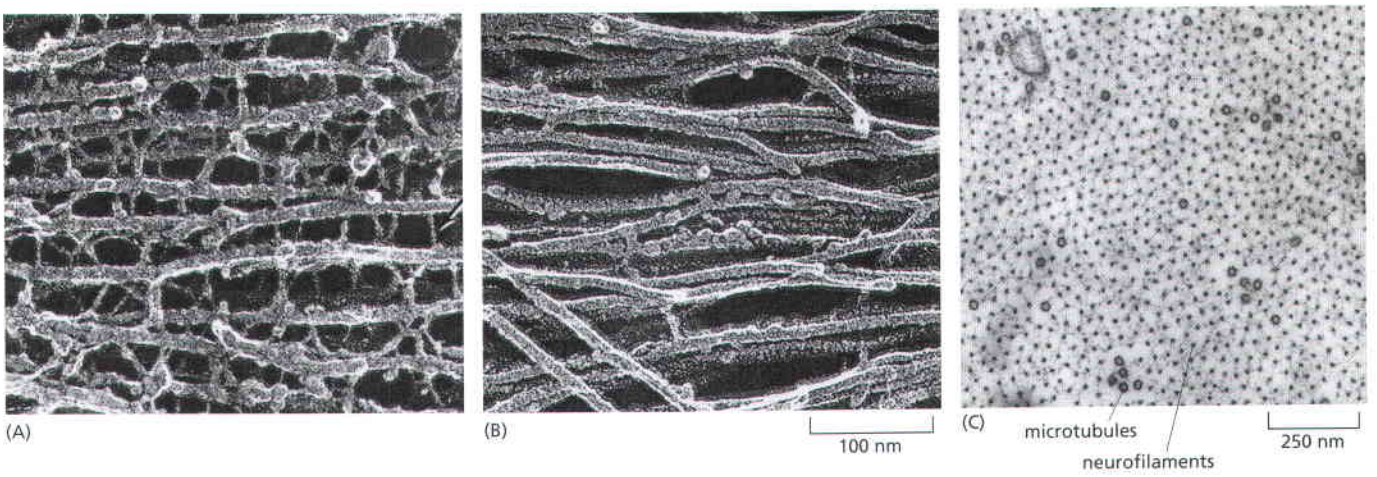


Figure 16–22 Two types of intermediate filaments in cells of the nervous system. (A) Freeze-etch electron microscopic image of neurofilaments in a nerve cell axon, showing the extensive cross-linking through protein cross-bridges—an arrangement believed to give this long cell process great tensile strength. The cross-bridges are formed by the long, nonhelical extensions at the C-terminus of the largest neurofilament protein (NF-H). (B) Freeze-etch image of glial filaments in glial cells, showing that these intermediate filaments are smooth and have few cross-bridges. (C) Conventional electron micrograph of a cross section of an axon showing the regular side-to-side spacing of the neurofilaments, which greatly outnumber the microtubules. (A and B, courtesy of Nobutaka Hirokawa; C, courtesy of John Hopkins.)

the central rod domain, demonstrating the importance of this particular part of the protein for correct filament assembly.

A second family of intermediate filaments, called **neurofilaments**, is found in high concentrations along the axons of vertebrate neurons (**Figure 16–22**). Three types of neurofilament proteins (NF-L, NF-M, NF-H) coassemble *in vivo*, forming heteropolymers that contain NF-L plus one of the others. The NF-H and NF-M proteins have lengthy C-terminal tail domains that bind to neighboring filaments, generating aligned arrays with a uniform interfilament spacing. During axonal growth, new neurofilament subunits are incorporated all along the axon in a dynamic process that involves the addition of subunits along the filament length, as well as the addition of subunits at the filament ends. After an axon has grown and connected with its target cell, the diameter of the axon may increase as much as fivefold. The level of neurofilament gene expression seems to directly control axonal diameter, which in turn influences how fast electrical signals travel down the axon.

The neurodegenerative disease amyotrophic lateral sclerosis (ALS, or Lou Gehrig's Disease) is associated with an accumulation and abnormal assembly of neurofilaments in motor neuron cell bodies and in the axon, which may interfere with normal axonal transport. The degeneration of the axons leads to muscle weakness and atrophy, which is usually fatal. The over-expression of human NF-L or NF-H in mice results in mice that have an ALS-like disease.

The vimentin-like filaments are a third family of intermediate filaments. Desmin, a member of this family, is expressed in skeletal, cardiac, and smooth muscle. Mice lacking desmin show normal initial muscle development, but adults have various muscle cell abnormalities, including misaligned muscle fibers.

Drugs Can Alter Filament Polymerization

Because the survival of eucaryotic cells depends on a balanced assembly and disassembly of the highly conserved cytoskeletal filaments formed from actin and tubulin, the two types of filaments are frequent targets for natural toxins. These toxins are produced in self-defense by plants, fungi, or sponges that do not wish to be eaten but cannot run away from predators, and they generally disrupt the filament polymerization reaction. The toxin binds tightly to either the filament form or the free subunit form of a polymer, driving the assembly reaction in the direction that favors the form to which the toxin binds. For example, the drug *latrunculin*, extracted from the sea sponge *Latrunculia magnifica*, binds to actin monomers and prevents their assembly into filaments; it thereby causes a net depolymerization of actin filaments. In contrast, *phalloidin*, from the fungus *Amanita phalloides* (death cap), binds to and stabilizes actin filaments, causing a net increase in actin polymerization. (This attractive but inedible mushroom also expresses a second deadly toxin, the RNA polymerase II inhibitor α-amanitin.) Either change in actin filaments is very toxic for cells. Similarly, *colchicine*, from the autumn crocus (or meadow saffron), binds to and

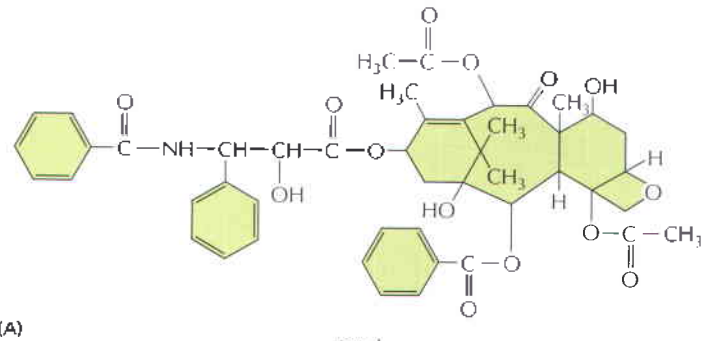
Table 16–2 **Drugs That Affect Actin Filaments and Microtubules**

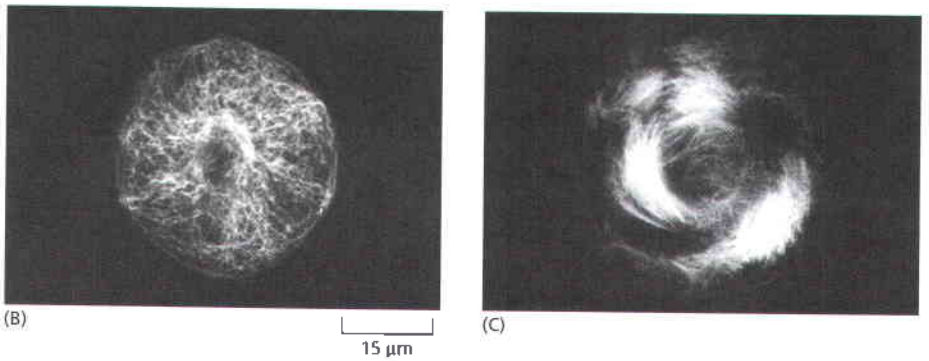
ACTIN-SPECIFIC DRUGS	
Phalloidin	binds and stabilizes filaments
Cytochalasin	caps filament plus ends
Swinholide	severs filaments
Latrunculin	binds subunits and prevents their polymerization
MICROTUBULE-SPECIFIC DRUGS	
Taxol	binds and stabilizes microtubules
Colchicine, colcemid	binds subunits and prevents their polymerization
Vinblastine, vincristine	binds subunits and prevents their polymerization
Nocodazole	binds subunits and prevents their polymerization

stabilizes free tubulin, causing microtubule depolymerization. In contrast, *taxol*, extracted from the bark of a rare species of yew tree, binds to and stabilizes microtubules, causing a net increase in tubulin polymerization. These and some other natural products that are commonly used by cell biologists to manipulate the cytoskeleton are listed in **Table 16–2**.

Drugs like these have a rapid and profound effect on the organization of the cytoskeleton in living cells (**Figure 16–23**). They provided early evidence that the cytoskeleton is a dynamic structure, maintained by a rapid and continual exchange of subunits between the soluble and filamentous forms, and they revealed that this subunit flux is necessary for normal cytoskeletal function.

The drugs listed in Table 16–2 have been useful to cell biologists trying to probe the roles of actin and microtubules in various cell processes. Some of them are also used to treat cancer. Both microtubule-depolymerizing drugs (such as vinblastine) and microtubule-polymerizing drugs (such as taxol) preferentially kill dividing cells, since both microtubule assembly and disassembly are crucial for correct function of the mitotic spindle (discussed later in this chapter). These drugs efficiently kill certain types of tumor cells in a human patient, although not without toxicity to rapidly dividing normal cells, including those in the bone marrow, intestine, and hair follicles. Taxol in particular has been widely used to treat cancers of the breast and lung, and it is frequently successful in treatment of tumors that are resistant to other chemotherapeutic agents.






(D)

Figure 16–23 Effect of the drug taxol on microtubule organization. (A) Molecular structure of taxol. Recently, organic chemists have succeeded in synthesizing this complex molecule, which is widely used for cancer treatment. (B) Immunofluorescence micrograph showing the microtubule organization in a liver epithelial cell before the addition of taxol. (C) Microtubule organization in the same type of cell after taxol treatment. Note the thick circumferential bundles of microtubules around the periphery of the cell. (D) A Pacific yew tree, the natural source of taxol. (B, C from N.A. Gloushankova et al., *Proc. Natl Acad. Sci. U.S.A.* 91:8597–8601, 1994. With permission from National Academy of Sciences; D, courtesy of A.K. Mitchell 2001. © Her Majesty the Queen in Right of Canada, Canadian Forest Service.)

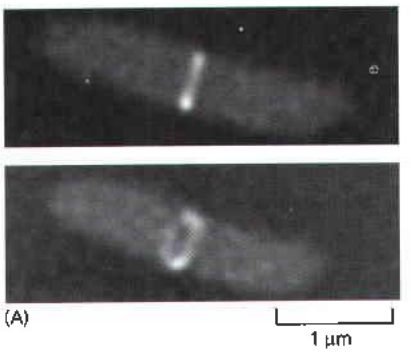
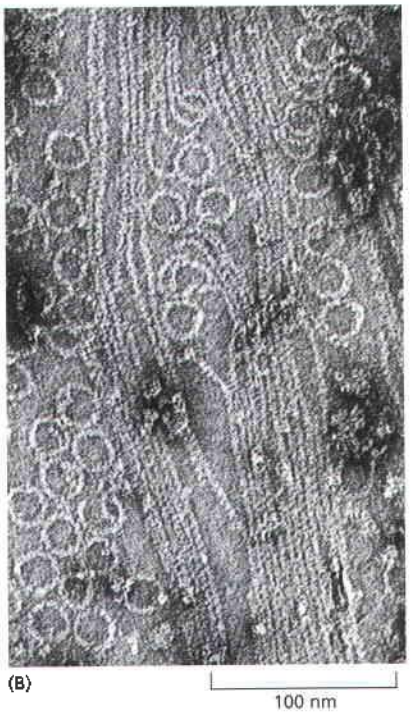


Figure 16–24 The bacterial FtsZ protein, a tubulin homolog in procaryotes. (A) A band of FtsZ protein forms a ring in a dividing bacterial cell. This ring has been labeled by fusing the FtsZ protein to the green fluorescent protein (GFP), which allows it to be observed in living *E. coli* cells with a fluorescence microscope. *Top*, side view shows the ring as a bar in the middle of the dividing cell. *Bottom*, rotated view showing the ring structure. (B) FtsZ filaments and rings, formed *in vitro*, as visualized using electron microscopy. Compare this image with that of the microtubule shown on the *right* in Figure 16–16C. (A, from X. Ma, D.W. Ehrhardt and W. Margolin, *Proc. Natl Acad. Sci. U.S.A.* 93:12998–13003, 1996; B, from H.A. Erickson et al., *Proc. Natl Acad. Sci. U.S.A.* 93:519–523, 1996. All with permission from National Academy of Sciences.)

Bacterial Cell Organization and Cell Division Depend on Homologs of the Eucaryotic Cytoskeleton



While eucaryotic cells are typically large and morphologically complex, bacterial cells are usually only a few micrometers long and assume simple, modest shapes such as spheres or rods. Bacteria also lack the elaborate networks of intracellular membrane-enclosed organelles such as the endoplasmic reticulum and Golgi apparatus. For many years, biologists assumed that the lack of a bacterial cytoskeleton was one reason for these striking differences between cell organization in the eucaryotic and bacterial kingdoms. This assumption was challenged with the discovery in the early 1990s that nearly all bacteria and many archaea contain a homolog of tubulin, FtsZ, that can polymerize into filaments and assemble into a ring (called the Z-ring) at the site where the septum forms during cell division (**Figure 16–24**).

The three-dimensional folded protein structure of FtsZ is remarkably similar to the structure of α or β tubulin and, like tubulin, hydrolysis of GTP is triggered by polymerization and causes a conformational change in the filament structure. Although the Z-ring itself persists for many minutes, the individual filaments within it are highly dynamic, with an average half-life of about thirty seconds. As the bacterium divides, the Z-ring becomes smaller until it has completely disassembled, and it is thought that the shrinkage of the Z-ring may contribute to the membrane invagination necessary for the completion of cell division. The Z-ring may also serve as a site for localization of specialized cell wall synthesis enzymes required for building the septum between the two daughter cells. The disassembled FtsZ subunits later reassemble at the new sites of septum formation in the daughter cells (**Figure 16–25**).

More recently, it has been found that many bacteria also contain homologs of actin. Two of these, MreB and Mbl, are found primarily in rod-shaped or spiral-shaped cells, and mutations disrupting their expression cause extreme abnormalities in cell shape and defects in chromosome segregation (**Figure 16–26**). MreB and Mbl filaments assemble *in vivo* to form large-scale spirals that

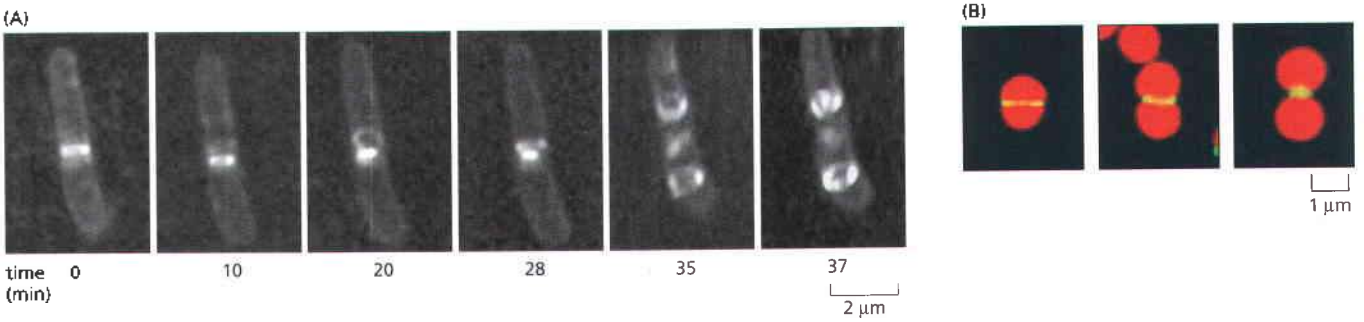


Figure 16–25 Rapid rearrangements of FtsZ through the bacterial cell cycle. (A) After chromosome segregation is complete, the ring formed by FtsZ at the middle of the cell becomes smaller as the cell pinches in two, much like the contractile ring formed by actin and myosin filaments in eucaryotic cells. The FtsZ filaments that have disassembled as the cells have separated then reassemble to form two new rings at the middle of the two daughter cells. (B) Dividing chloroplasts (*red*) from a red alga also make use of a protein ring made from FtsZ (*yellow*) for cleavage. (A, from Q. Sun and W. Margolin, *J. Bacteriol.* 180:2050–2056, 1998. With permission from American Society for Microbiology; B, from S. Miyagishima et al., *Plant Cell* 13:2257–2268, 2001. With permission from American Society of Plant Biologists.)

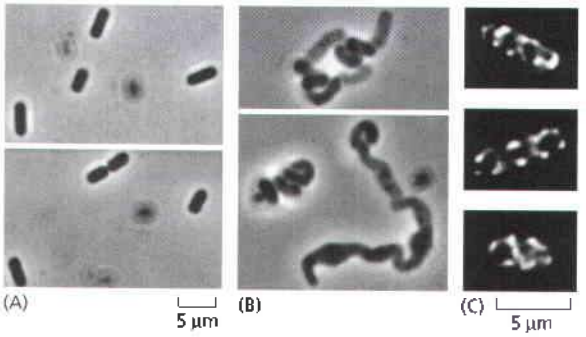


Figure 16–26 Actin homologs in bacteria determine cell shape. (A) The common soil bacterium *Bacillus subtilis* normally forms cells with a regular rod-like shape. (B) *B. subtilis* cells lacking the actin homolog Mbl grow into irregular twisted tubes and eventually die. (C) The Mbl protein forms long helices made of up many short filaments that run the length of the bacterial cell and help to direct the sites of cell wall synthesis. (From L.J. Jones, R. Carbadillo-Lopez and J. Errington, *Cell* 104: 913-922, 2001. With permission from Elsevier.)

span the length of the cell and apparently contribute to cell shape determination by serving as a scaffold to direct the synthesis of the peptidoglycan cell wall, in much the same way that microtubules help organize the synthesis of the cellulose cell wall in higher plant cells (see Figure 19–82). As with FtsZ, the filaments within the MreB and Mbl spirals are highly dynamic, with half-lives of a few minutes; as for actin, ATP hydrolysis accompanies the polymerization process.

Diverse relatives of MreB and Mbl have more specialized roles. A particularly intriguing bacterial actin homolog is ParM, which is encoded on certain bacterial plasmids that also carry genes responsible for antibiotic resistance and frequently cause the spread of multi-drug resistance in epidemics. Bacterial plasmids typically encode all the gene products that are necessary for their own segregation, presumably as a strategy to ensure their faithful inheritance and propagation in their bacterial hosts. *In vivo*, ParM assembles into a filamentous structure that associates at each end with a copy of the plasmid that encodes it, and growth of the ParM filament appears to push the replicated plasmid copies apart, rather like a mitotic spindle operating in reverse (**Figure 16–27**). Although ParM is a structural homolog of actin, its dynamic behavior differs significantly. ParM filaments undergo dramatic dynamic instability *in vitro*, more closely resembling microtubules than actin filaments in the way that they grow and shrink. The spindle-like structure is apparently built by the selective stabilization of spontaneously nucleated filaments that bind to specialized proteins recruited to the origins of replication on the plasmids.

The various bacterial actin homologs share similar molecular structures but their amino acid sequence similarity to each other is quite low (~10–15% identical residues). They assemble into filaments with distinct helical packing patterns, which may also have very different dynamic behaviors. Rather than using the same well-conserved actin for many different purposes, as eucaryotic cells do, bacteria have apparently opted to proliferate and specialize their actin homologs for distinct purposes.

It is now clear that the general principle of organizing cell structure by the self-association of nucleotide-binding proteins into dynamic helical filaments is used in all cells, and that the two major families of actin and tubulin are very

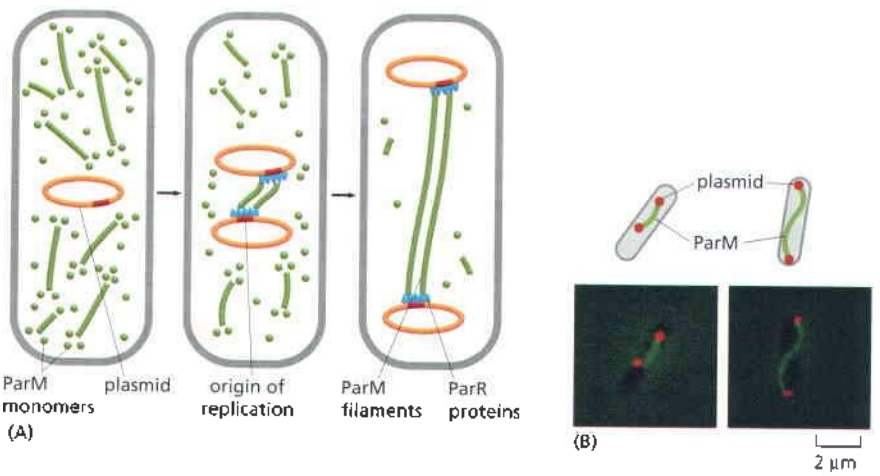


Figure 16–27 Role of the actin homolog ParM in plasmid segregation. (A) Some bacterial drug-resistance plasmids (*yellow*) encode an actin homolog, ParM, that will spontaneously nucleate to form small, dynamic filaments (*green*) throughout the bacterial cytoplasm. A second plasmid-encoded protein (*blue*) binds to specific DNA sequences in the plasmid, and also stabilizes the dynamic ends of the ParM filaments. When the plasmid has duplicated, so that the ParM filaments can be stabilized at both ends, the filaments grow and push the duplicated plasmids to opposite ends of the cell. (B) In these bacterial cells harboring a drug-resistance plasmid, the plasmids are labeled in *red* and the ParM protein in *green*. Left, a short ParM bundle connects the two daughter plasmids shortly after their duplication. Right, the fully assembled ParM filament has pushed the duplicated plasmids to the cell poles. (A, adapted from E.C. Garner, C.S. Campbell and R.D. Mullins, *Science* 306:1021–1025, 2004. With permission from AAAS; B, from J. Moller-Jensen et al., *Mol. Cell* 12:1477–1487, 2003. With permission from Elsevier.)

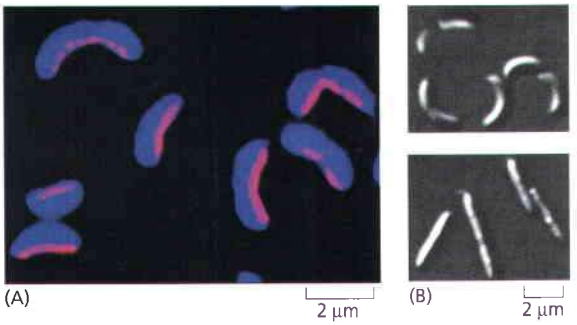


Figure 16–28 Caulobacter and crescentin. The sickle-shaped bacterium *Caulobacter crescentus* expresses a protein, crescentin, with a series of coiled-coil domains similar in size and organization to the domains of eucaryotic intermediate filaments. In cells, the crescentin protein forms a fiber that runs down the inner side of the curving bacterial cell wall. When the gene is disrupted, the bacteria are viable but grow in a straight rod-shaped form. (From N. Ausmees, J.R. Kuhn and C. Jacobs-Wagner, *Cell* 115:705–713, 2003. With permission from Elsevier.)

ancient, probably predating the split between the eucaryotic and bacterial kingdoms. However, the uses to which bacteria put their cytoskeletons appear somewhat different from their eucaryotic homologs. For example, in bacteria it is the tubulin (FtsZ) that is involved in *cytokinesis* (the pinching apart of a dividing cell into two daughters), while actin drives this process in eucaryotic cells. Conversely, eucaryotic microtubules are responsible for chromosome segregation, while bacterial actins (ParM and possibly MreB) help to segregate replicated DNA in bacteria.

At least one bacterial species with an unusual crescent shape, *Caulobacter crescentus*, even appears to harbor a protein with significant structural similarity to the third major class of cytoskeletal filaments found in animal cells, the intermediate filaments. A protein called crescentin forms a filamentous structure that seems to influence the cell shape; when the gene encoding crescentin is deleted, the *Caulobacter* cells grow as straight rods (**Figure 16–28**).

Since we now know that bacteria do in fact have sophisticated dynamic cytoskeletons, why then do they remain so small and morphologically simple? As yet there have been no *motor proteins* identified that walk along the bacterial filaments; perhaps the evolution of motor proteins was a critical step allowing morphological elaboration in the eucaryotes.

Summary

The cytoplasm of eucaryotic cells is spatially organized by a network of protein filaments known as the cytoskeleton. This network contains three principal types of filaments: microtubules, actin filaments, and intermediate filaments. All three types of filaments form as helical assemblies of subunits that self-associate using a combination of end-to-end and side-to-side protein contacts. Differences in the structure of the subunits and the manner of their self-assembly give the filaments different mechanical properties. Intermediate filaments are rope-like and easy to bend but hard to break. Microtubules are strong, rigid hollow tubes. Actin filaments are the thinnest of the three and are easy to break.

In living cells, the assembly and disassembly of their subunits constantly remodels all three types of cytoskeletal filaments. Microtubules and actin filaments add and lose subunits only at their ends, with one end (the plus end) growing faster than the other. Tubulin and actin (the subunits of microtubules and actin filaments, respectively) bind and hydrolyze nucleoside triphosphates (tubulin binds GTP and actin binds ATP). Nucleotide hydrolysis underlies the characteristic dynamic behavior of these two filaments. Actin filaments in cells seem to predominantly undergo treadmilling, where a filament assembles at one end while simultaneously disassembling at the other end. Microtubules in cells predominantly display dynamic instability, where a microtubule end undergoes alternating bouts of growth and shrinkage.

Whereas tubulin and actin have been strongly conserved in evolution, the family of intermediate filaments is very diverse. There are many tissue-specific forms found in the cytoplasm of animal cells, including keratin filaments in epithelial cells, neurofilaments in nerve cells, and desmin filaments in muscle cells. In all these cells, the primary job of intermediate filaments is to provide mechanical strength.

Bacterial cells also contain homologs of tubulin, actin and intermediate filaments that form dynamic filamentous structures involved in determining cell shape and in cell division.

HOW CELLS REGULATE THEIR CYTOSKELETAL FILAMENTS

Microtubules, actin filaments, and intermediate filaments are much more dynamic in cells than they are in the test tube. The cell regulates the length and stability of its cytoskeletal filaments, as well as their number and geometry. It does so largely by regulating their attachments to one another and to other components of the cell, so that the filaments can form a wide variety of higher-order structures. Direct covalent modification of the filament subunits regulates some filament properties, but most of the regulation is performed by a large array of accessory proteins that bind to either the filaments or their free subunits. Some of the most important accessory proteins associated with microtubules and actin filaments are outlined in **Panel 16–3** (pp. 994–995). This section describes how these accessory proteins modify the dynamics and structure of cytoskeletal filaments. We begin with a discussion of the way that microtubules and actin filaments are nucleated in cells, because this plays a major part in determining the overall organization of the cell's interior.

A Protein Complex Containing γ-Tubulin Nucleates Microtubules

While α- and β-tubulins are the regular building blocks of microtubules, another type of tubulin, called *γ-tubulin*, has a more specialized role. Present in much smaller amounts than α- and β-tubulin, this protein is involved in the nucleation of microtubule growth in organisms ranging from yeasts to humans. Microtubules are generally nucleated from a specific intracellular location known as a **microtubule-organizing center** (**MTOC**). Antibodies against γ-tubulin stain the MTOC in virtually all species and cell types thus far examined.

Microtubules are nucleated at their minus end, with the plus end growing outward from each MTOC to create various types of microtubule arrays. A **γ-tubulin ring complex** (**γ-TuRC**) that is capable of nucleating microtubule growth in a test tube has been isolated from both insect and vertebrate cells. Two proteins, conserved from yeasts to humans, bind directly to the γ-tubulin, along with several other proteins that help create a ring of γ-tubulin molecules. This ring can be seen at the minus ends of the microtubules nucleated by γ-TuRC, and it is therefore thought to serve as a template that creates a microtubule with 13 protofilaments (**Figure 16–29**).

Microtubules Emanate from the Centrosome in Animal Cells

Most animal cells have a single, well-defined MTOC called the **centrosome**, located near the nucleus. From this focal point, the cytoplasmic microtubules emanate in a star-like, "astral" conformation. Microtubules are nucleated at the centrosome at their minus ends, so the plus ends point outward and grow toward the cell periphery. Microtubules nucleated at the centrosome continuously grow and shrink by dynamic instability, probing the entire three-dimensional volume of the cell. A centrosome is composed of a fibrous *centrosome*

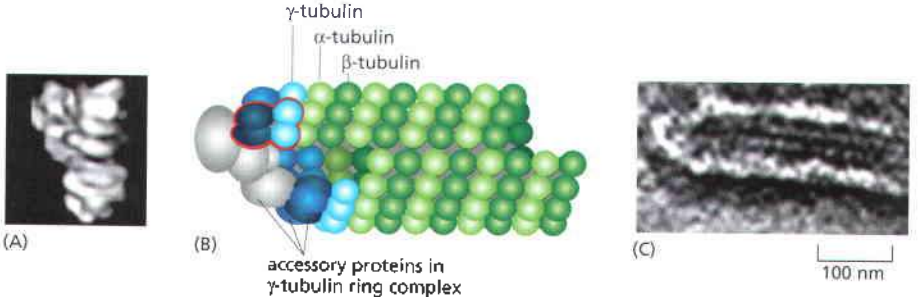
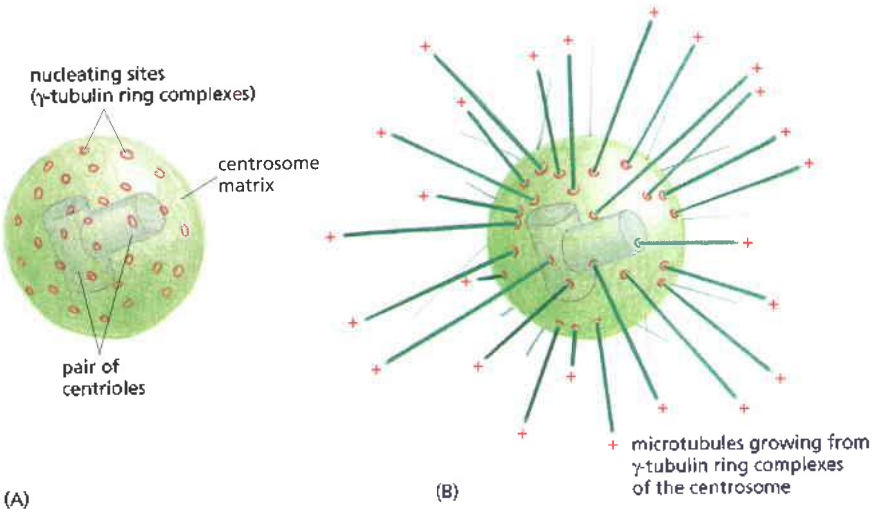


Figure 16–29 Polymerization of tubulin nucleated by γ-tubulin ring complexes. (A) Structure of the γ-tubulin ring complex, reconstructed from averaging electron micrographs of individual purified complexes. (B) Model for the nucleation of microtubule growth by the γ-TuRC. The red outline indicates a pair of proteins bound to two molecules of γ-tubulin; this group can be isolated as a separate subcomplex of the larger ring. Note the longitudinal discontinuity between two protofilaments. Microtubules generally have one such "seam" breaking the otherwise uniform helical packing of the protofilaments. (C) Electron micrograph of a single microtubule nucleated from the purified γ-tubulin ring complex. (A and C, from M. Moritz et al., *Nat. Cell Biol.* 2:365–370, 2000. With permission from Macmillan Publishers Ltd.)



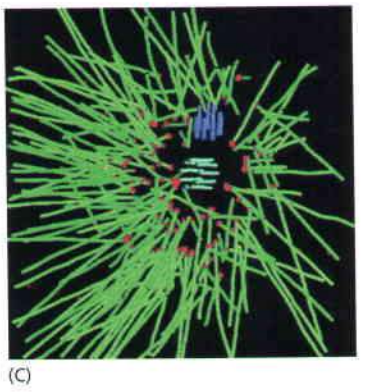

(C)

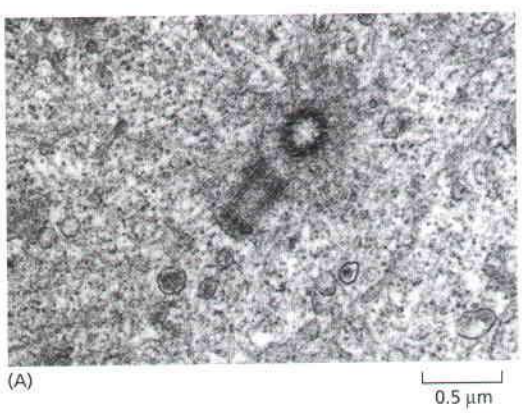
Figure 16–30 The centrosome. (A) The centrosome is the major MTOC of animal cells. Located in the cytoplasm next to the nucleus, it consists of an amorphous matrix of fibrous proteins to which the γ-tubulin ring complexes that nucleate microtubule growth are attached. This matrix is organized by a pair of centrioles, as described in the text. (B) A centrosome with attached microtubules. The minus end of each microtubule is embedded in the centrosome, having grown from a γ-tubulin ring complex, whereas the plus end of each microtubule is free in the cytoplasm. (C) In a reconstructed image of the MTOC from a *C. elegans* cell, a dense thicket of microtubules can be seen emanating from the centrosome. (C, from E.T. O'Toole et al., *J. Cell Biol.* 163:451–456, 2003. With permission from The Rockefeller University Press.)

matrix that contains more than fifty copies of γ-TuRC. Most of the proteins that form this matrix remain to be discovered, and it is not yet known how they recruit and activate the γ-TuRC.

Embedded in the centrosome is a pair of somewhat mysterious cylindrical structures arranged at right angles to each other in an L-shaped configuration (Figure 16–30). These are the **centrioles**, which become the basal bodies of cilia and flagella in motile cells (described later). The centrioles organize the centrosome matrix (also called the pericentriolar material), ensuring its duplication during each cell cycle as the centrioles themselves duplicate (Figure 16–31). As described in Chapter 17, the centrosome duplicates and splits into two equal parts during interphase, each half containing a duplicated centriole pair. These two daughter centrosomes move to opposite sides of the nucleus when mitosis begins, and they form the two poles of the mitotic spindle (see Panel 17–1, pp. 1072–1073). A centriole consists of a short cylinder of modified microtubules, plus a large number of accessory proteins. The molecular basis for its duplication is not well-understood.

In fungi and diatoms, microtubules are nucleated at an MTOC that is embedded in the nuclear envelope as a small plaque called the *spindle pole body*. Higher-plant cells seem to nucleate microtubules at sites distributed all around the nuclear envelope. Neither fungi nor most plant cells contain centrioles. Despite these differences, all these cells contain γ-tubulin and seem to use it to nucleate their microtubules.

In animal cells, the astral configuration of microtubules is very robust, with dynamic plus ends pointing outward toward the cell periphery and stable minus ends collected near the nucleus. The system of microtubules radiating from the centrosome acts as a device to survey the outlying regions of the cell and to position the centrosome at its center, and it does this even in artificial enclosures



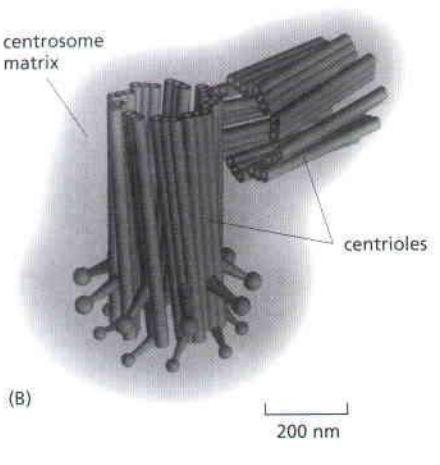
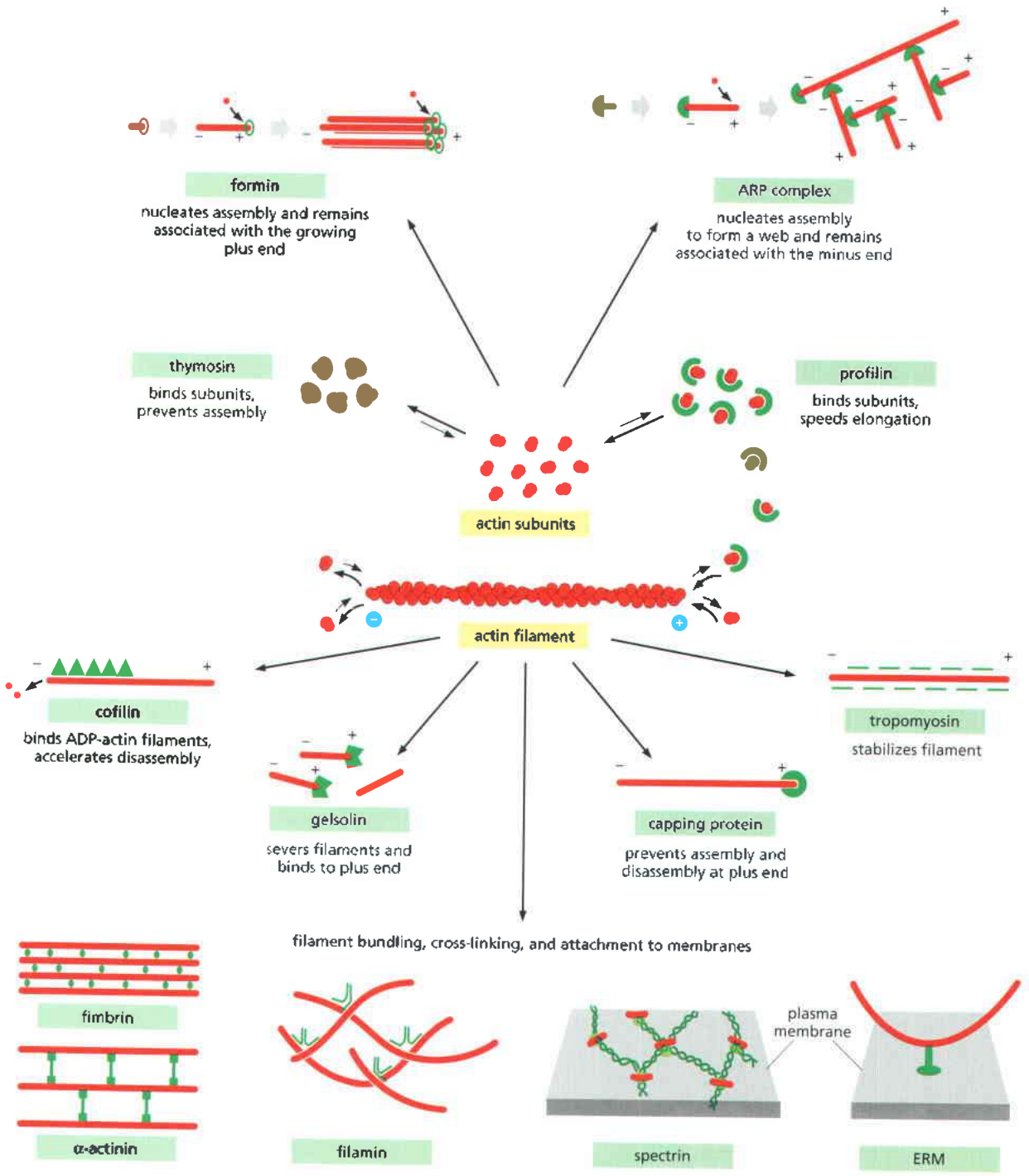


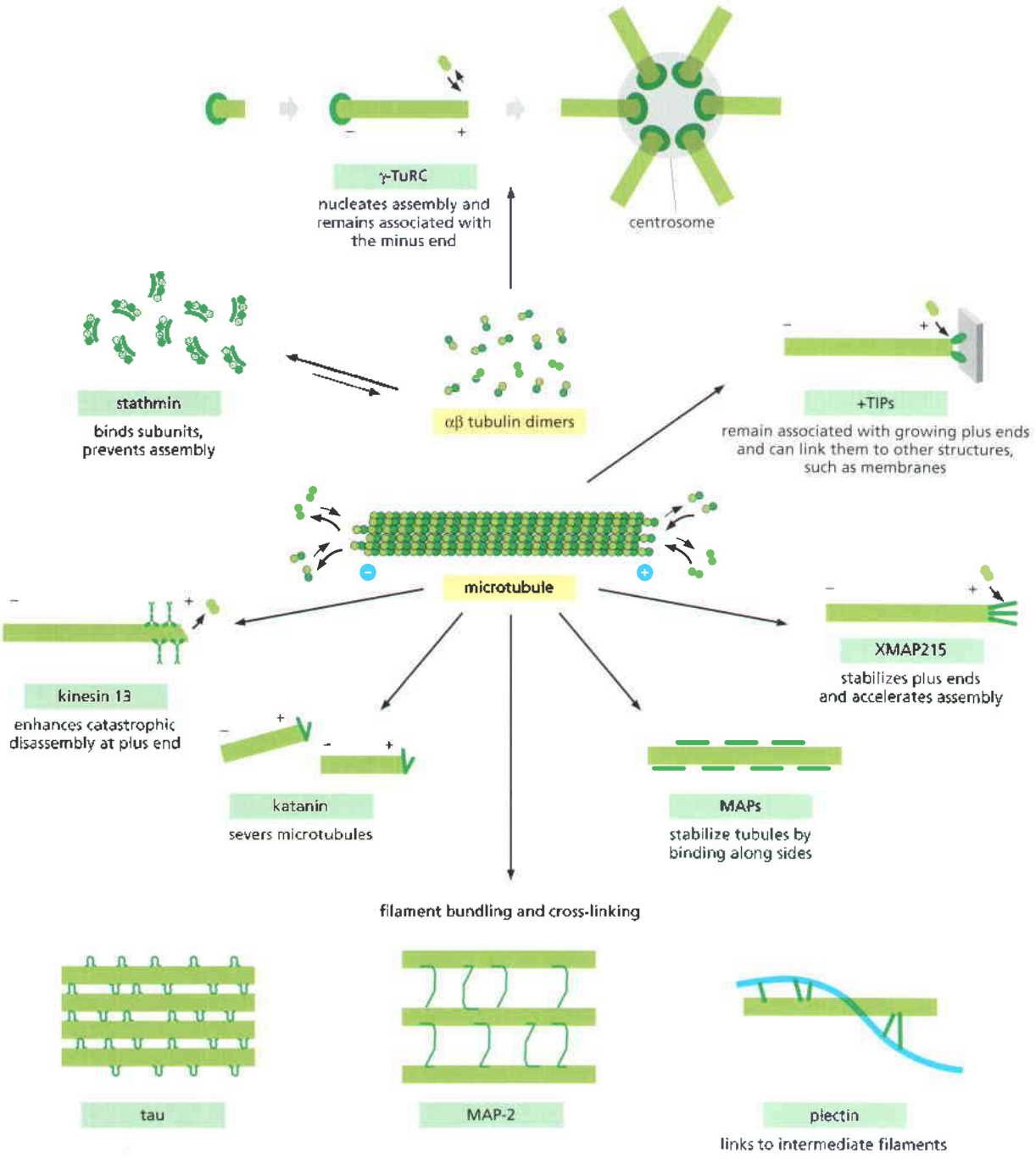
Figure 16–31 A centriole in the centrosome. (A) An electron micrograph of a thin section of a centrosome showing an end-on view of the mother centriole and a longitudinal section of the daughter centriole. Numerous microtubules are seen nearby. (B) Structure of the centriole pair. (A, from G.J. Mack, Y. Ou and J.B. Rattner, *Microsc. Res. Tech.* 49:409–419, 2000. With permission from John Wiley & Sons. B, adapted from D. Chrétien et al., *J. Struct. Biol.* 120:117–133, 1997. With permission from Elsevier.)

ACTIN FILAMENTS



Some of the major accessory proteins of the actin cytoskeleton. Except for the myosin motor proteins, to be discussed in a later section, an example of each major type is shown. Each of these is discussed in the text. However, most cells contain more than a hundred different actin-binding proteins, and it is likely that there are important types of actin-associated proteins that are not yet recognized.

MICROTUBULES



Some of the major accessory proteins of the microtubule cytoskeleton. Except for two classes of motor proteins, to be discussed in a later section, an example of each major type is shown. Each of these is discussed in the text. However, most cells contain more than a hundred different microtubule-binding proteins, and—as for the actin-associated proteins—it is likely that there are important types of microtubule-associated proteins that are not yet recognized.

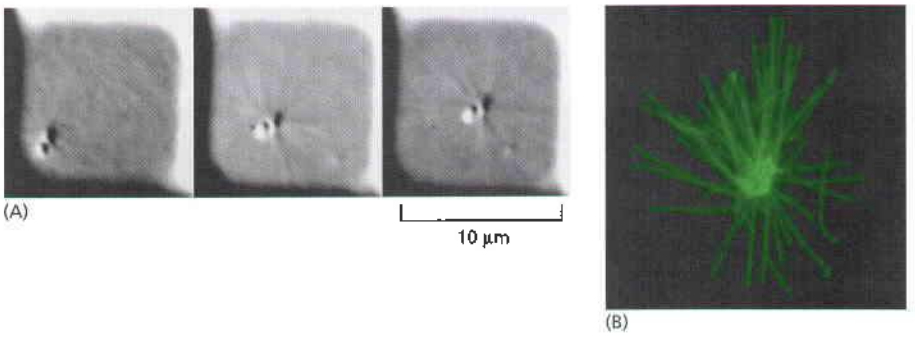


Figure 16–32 The center-seeking behavior of a centrosome. (A) Small square wells were micromachined into a plastic substrate. A single centrosome was placed into one of these wells, along with tubulin subunits in solution. As the microtubules polymerize, nucleated by the centrosome, they push against the walls of the well. The requirement for equal pushing in all directions to stabilize the position forces the centrosome to the center of the well. These pictures were taken at three-minute intervals. (B) A self-centered centrosome, fixed and stained to show the distribution of the microtubules pushing on all four walls of the enclosure. (From T.E. Holy et al., *Proc. Natl Acad. Sci. U.S.A.* 94:6228–6231, 1997. With permission from National Academy of Sciences.)

(**Figure 16–32**). Even in an isolated cell fragment lacking the centrosome, dynamic microtubules interacting with membraneous organelles arrange themselves into a star-shaped array with the microtubule minus ends clustered at the center, although this process may involve more components than just the simple pushing mechanism used by the isolated centrosome (**Figure 16–33**). This ability of the microtubule cytoskeleton to find the center of the cell establishes a general coordinate system, which is then used to position many organelles within the cell. Highly differentiated cells with complex morphologies such as neurons, muscles, and epithelial cells must use additional measuring mechanisms to establish their more elaborate internal coordinate systems. Thus, for example, when an epithelial cell forms cell–cell junctions and becomes highly polarized, the microtubule minus ends move to a region near the apical plasma membrane. From this asymmetric location, an array of nearly parallel microtubules forms along the long axis of the cell, with plus ends extending as far as the basal surface (see Figure 16–5).

Actin Filaments Are Often Nucleated at the Plasma Membrane

In contrast to microtubule nucleation, which occurs primarily deep within the cytoplasm near the nucleus, actin filament nucleation most frequently occurs at or near the plasma membrane. Consequently, the highest density of actin filaments in most cells is at the cell periphery. The layer just beneath the plasma membrane is called the **cell cortex**, and the actin filaments in it determine the shape and movement of the cell surface. For example, depending on their attachments to one another and to the plasma membrane, actin structures can form many strikingly different types of cell surface projections. These include spiky bundles such as *microvilli* or *filopodia*, flat protrusive veils called *lamellipodia* that help move cells over solid substrates, and the phagocytic cups in macrophages.

External signals frequently regulate the nucleation of actin filaments at the plasma membrane, allowing the cell to change its shape and stiffness rapidly in response to changes in its external environment. This nucleation can be catalyzed

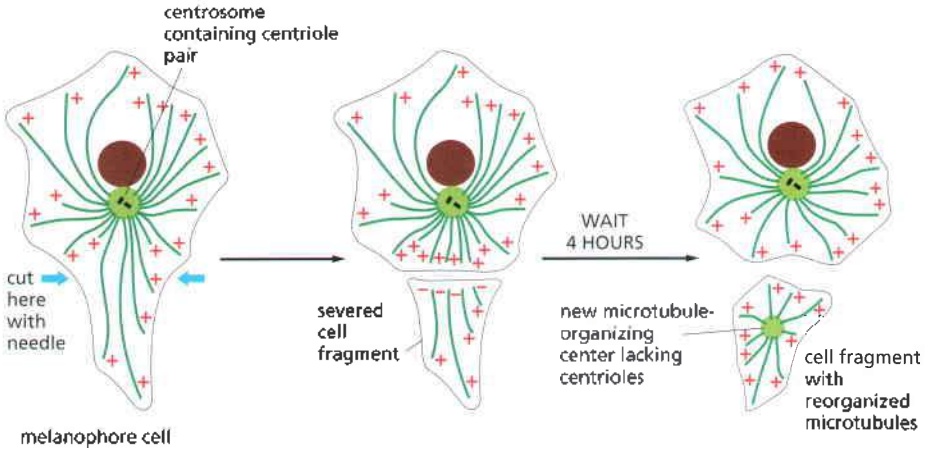


Figure 16–33 A microtubule array can find the center of a cell. After the arm of a fish pigment cell is cut off with a needle, the microtubules in the detached cell fragment reorganize so that their minus ends end up near the center of the fragment, buried in a new microtubule-organizing center.

by two different types of regulated factors, the ARP complex and the formins (discussed below). The first of these is a complex of proteins that includes two *actin-related proteins*, or *ARPs*, each of which is about 45% identical to actin. Analogous to the function of the γ-TuRC, the **ARP complex** (also known as the *Arp 2/3 complex*) nucleates actin filament growth from the minus end, allowing rapid elongation at the plus end (**Figure 16–34**A and B). The complex can also

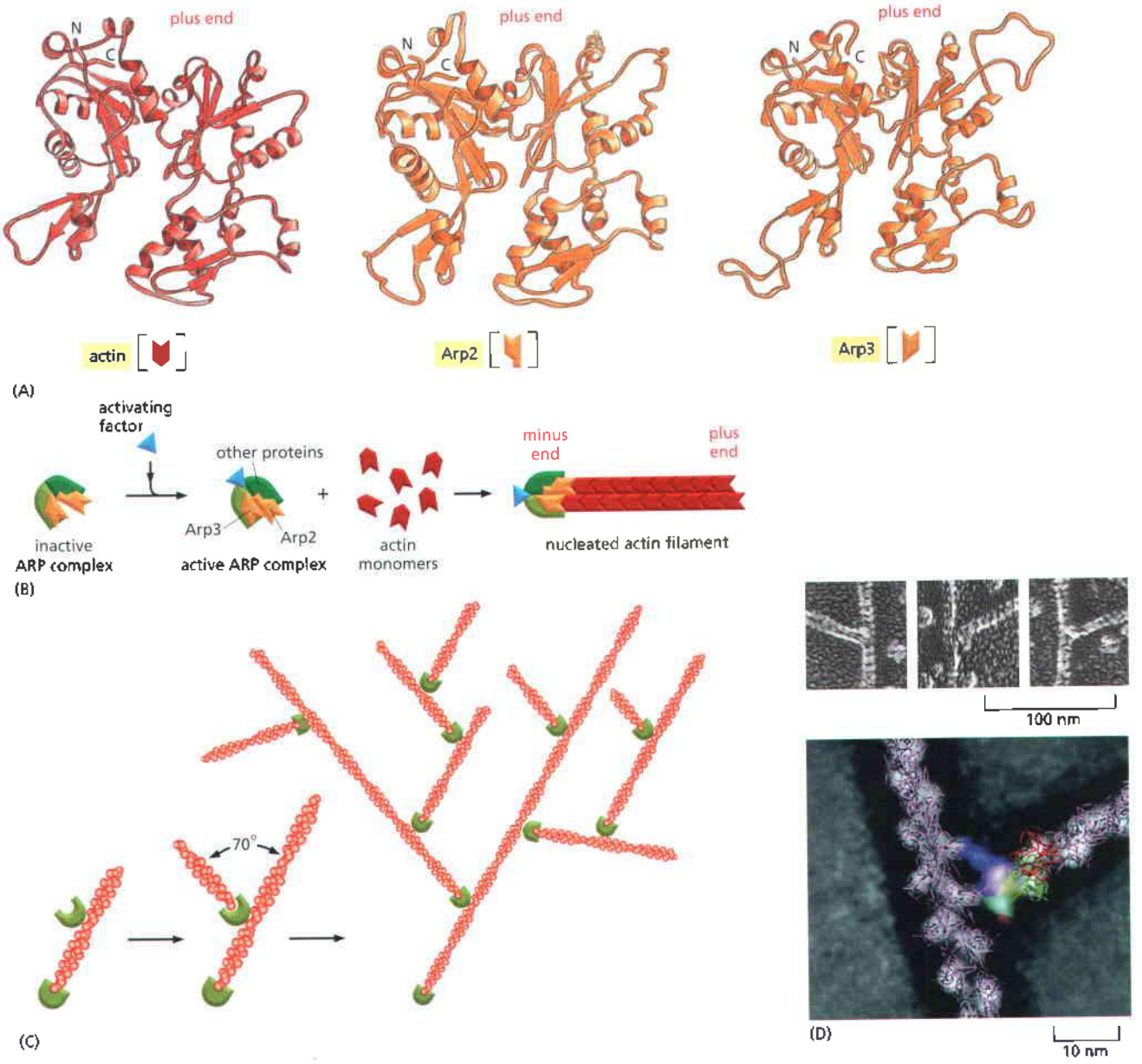


Figure 16–34 Nucleation and actin web formation by the ARP complex. (A)The structures of Arp2 and Arp3, compared to the structure of actin. Although the face of the molecule equivalent to the plus end *(top)* in both Arp2 and Arp3 is very similar to the plus end of actin itself, differences on the sides and minus end *(bottom)* prevent these actin-related proteins from forming filaments on their own or coassembling into filaments with actin. (B) A model for actin filament nucleation by the ARP complex. In the absence of an activating factor, Arp2 and Arp3 are held by their accessory proteins in an orientation that prevents them from nucleating a new actin filament. When an activating factor indicated by the *blue triangle* binds the complex, Arp2 and Arp3 are brought together into a new configuration that resembles the plus end of an actin filament. Actin subunits can then assemble onto this structure, bypassing the rate-limiting step of filament nucleation (see Figure 16–10). (C) The ARP complex nucleates filaments most efficiently when it is bound to the side of a preexisting actin filament. The result is a filament branch that grows at a 70° angle relative to the original filament. Repeated rounds of branching nucleation result in a treelike web of actin filaments. (D) Top, electron micrographs of branched actin filaments formed by mixing purified actin subunits with purified ARP complexes. Bottom, reconstructed image of a branch where the crystal structures of actin and the ARP complex have been fitted to the electron density. The mother filament runs from top to bottom, and the daughter filament branches off to the right where the ARP complex binds to three actin subunits in the mother filament (D, from R.D. Mullins et al., *Proc. Natl Acad. Sci. U.S.A.* 95:6181–6186, 1998. With permission from National Academy of Sciences, and from N. Volkmann et al., *Science* 293:2456–2459, 2001. With permission from Macmillan Publishers Ltd.)

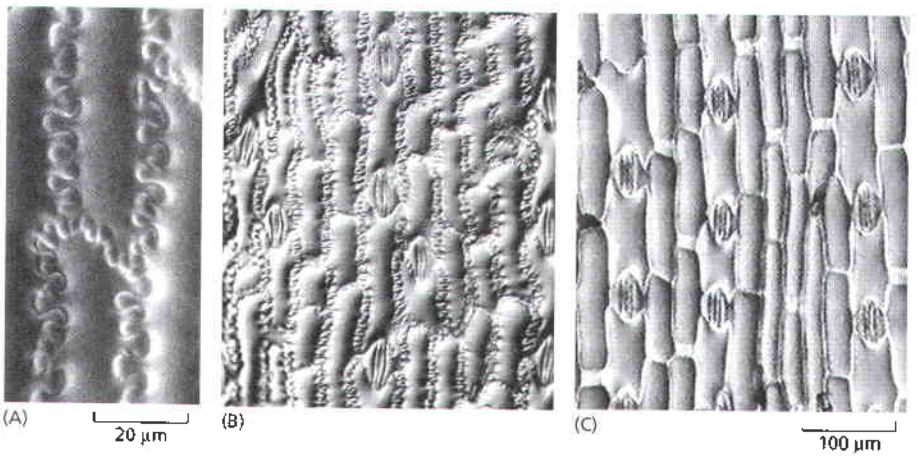


Figure 16–35 Function of the ARP complex in plant cells. (A) Cells in the maize leaf epidermis form small, actin-rich lobes that lock neighboring cells together like pieces of a jigsaw puzzle. (B) The regular pattern of interlocking cells covers the leaf surface. (C) Epidermal cells in a mutant plant lacking the ARP complex do not form the interlocking lobes. The brick-shaped cells are normal in size and spacing, but form leaves that appear too shiny to the naked eye. (From M.J. Frank, H.N. Cartwright and L.G. Smith, *Development* 130:753–762, 2003. With permission from the Company of Biologists.)

attach to the side of another actin filament while remaining bound to the minus end of the filament that it has nucleated, thereby building individual filaments into a treelike web (Figure 16–34C and D).

In animals, the ARP complex is associated with structures at the leading edge of migrating cells. The complex is localized in regions of rapid actin filament growth such as lamellipodia, and intracellular signaling molecules and components at the cytosolic face of the plasma membrane regulate its nucleating activity. This conserved complex is also involved in actin filament nucleation near the plasma membrane in yeast, where it is required to form cortical actin patches (see Figure 16–6), and in plant cells, where it directs the formation of actin bundles at the surface that are required for the growth of complex cell shapes in a variety of different tissues (**Figure 16–35**).

Both γ-tubulin and ARPs are evolutionarily ancient, and they are conserved among a wide variety of eucaryotic species. Their genes seem to have arisen by early duplication of the gene for the microtubule or actin filament subunit, respectively, followed by divergence and specialization of the gene copies so that they encode proteins with a special nucleating function. Thus, a similar strategy has evolved for two separate cytoskeletal systems. This underlines the central importance of regulated filament nucleation as a general organizing principle in cells.

The Mechanism of Nucleation Influences Large-Scale Filament Organization

Because the ARP complex nucleates the growth of a new actin filament most efficiently when it is bound to the side of an old actin filament, regulated activation of the ARP complex in animal cells tends to lead to the assembly of gel-like branched actin networks. However, many of the large-scale actin structures seen in cells are made up of parallel bundles of unbranched actin filaments, including the cleavage furrow found at the equator of dividing cells (see Figure 16–2) and the actin cables that point toward the site of bud growth in yeast (see Figure 16–6). The formation of many of these actin bundles is induced by a different set of nucleating proteins, the *formins*, which are able to nucleate the growth of straight, unbranched filaments that can be cross-linked by other proteins to form parallel bundles.

Formins are a large family of dimeric proteins (about 15 distinct formins are encoded in the mouse genome). Each formin subunit has a binding site for monomeric actin, and the formin dimer appears to nucleate actin filament polymerization by capturing two monomers. As the newly nucleated filament grows, the formin dimer remains associated with the rapidly growing plus end, while still allowing the binding of new subunits at that end to elongate the filament (**Figure 16–36**). This is very different from the behavior of the ARP complex or the γ-TuRC, which remain stably bound to the minus end of the actin filament or microtubule and prevent both subunit addition and subunit loss at this end.

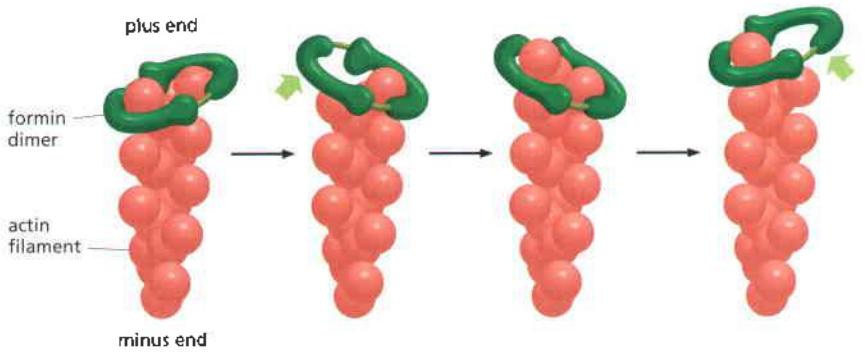


Figure 16–36 Actin elongation mediated by formins. Formin proteins (*green*) form a dimeric complex that can nucleate the formation of a new actin filament (*red*) and remain associated with the rapidly-growing plus end as it elongates. The formin protein maintains its binding to one of the two actin subunits exposed at the plus end as it allows each new subunit to assemble. Only part of the large formin molecule is shown here. Other regions regulate its activity and link it to particular structures in the cell. Many formins are indirectly connected to the cell plasma membrane, and aid the insertional polymerization of the actin filament directly beneath the membrane surface.

Proteins That Bind to the Free Subunits Modify Filament Elongation

Once nucleated, cytoskeletal filaments generally elongate by the addition of soluble subunits. In most nonmuscle vertebrate cells, approximately 50% of the actin is in filaments and 50% is soluble, although this ratio can change rapidly in response to external signals. The soluble monomer concentration is typically 50–200 μM (2–8 mg/ml); this is surprisingly high, given the critical concentration of less than 1 μM observed for pure actin in a test tube. Why does so much of the actin remain soluble, rather than polymerizing into filaments? The reason is that the subunit pool contains special proteins that bind to the actin monomers, thereby making polymerization much less favorable (an action similar to that of the drug latrunculin). A small protein called *thymosin* is the most abundant of these proteins. Actin monomers bound to thymosin are in a locked state, where they cannot associate with either the plus or minus ends of actin filament and can neither hydrolyze nor exchange their bound nucleotide.

How do cells recruit actin monomers from this buffered storage pool and use them for polymerization? It might seem as if signal transduction pathways such as those discussed in Chapter 15 could regulate thymosin, but this has not been found to be the case. Instead, recruitment depends on another monomer-binding protein, *profilin*. Profilin binds to the face of the actin monomer opposite the ATP-binding cleft, blocking the side of the monomer that would normally associate with the filament minus end, while leaving exposed the site on the monomer that binds to the plus end (**Figure 16–37**). The profilin–actin complex can readily add onto a free plus end. This addition induces a conformational change in the actin that reduces its affinity for profilin, so the profilin falls off, leaving the actin filament one subunit longer. Because profilin competes with thymosin in binding to individual actin monomers, the net result of a local activation of profilin molecules is a movement of actin subunits from the sequestered thymosin-bound pool onto filament plus ends. Actin filament growth depends even more strongly on profilin activation for those filaments whose plus end is associated with certain formins (the family of actin-nucleating proteins discussed above); in these cases, actin filament elongation can require that the monomeric actin be bound to profilin (**Figure 16–38**).

Several intracellular mechanisms regulate profilin activity, including profilin phosphorylation and profilin binding to inositol phospholipids. These mechanisms can define the sites where profilin acts. For example, profilin's ability to move sequestered actin subunits onto the growing ends of filaments is critical for filament assembly at the plasma membrane. Profilin is localized at the cytosolic face of the plasma membrane because it binds to acidic membrane phospholipids there. At this location, extracellular signals can activate profilin to produce explosive local actin polymerization and the extension of actin-rich

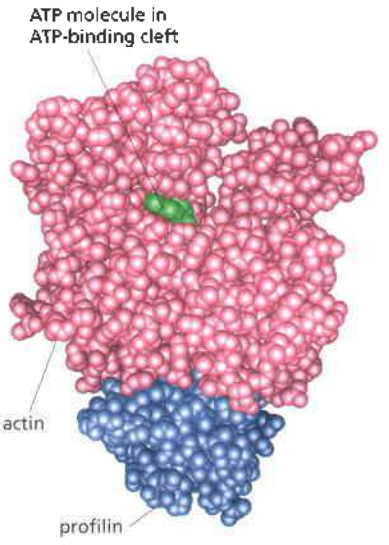


Figure 16–37 Profilin bound to an actin monomer. The profilin protein molecule is shown in *blue*, and the actin in *red*. ATP is shown in *green*. Profilin binds to the face of actin opposite the ATP-binding cleft. This profilin–actin heterodimer can therefore bind to and elongate the plus end of an actin filament, but it is sterically prevented from binding to the minus end. (Courtesy of Michael Rozycki and Clarence E. Schutt.)

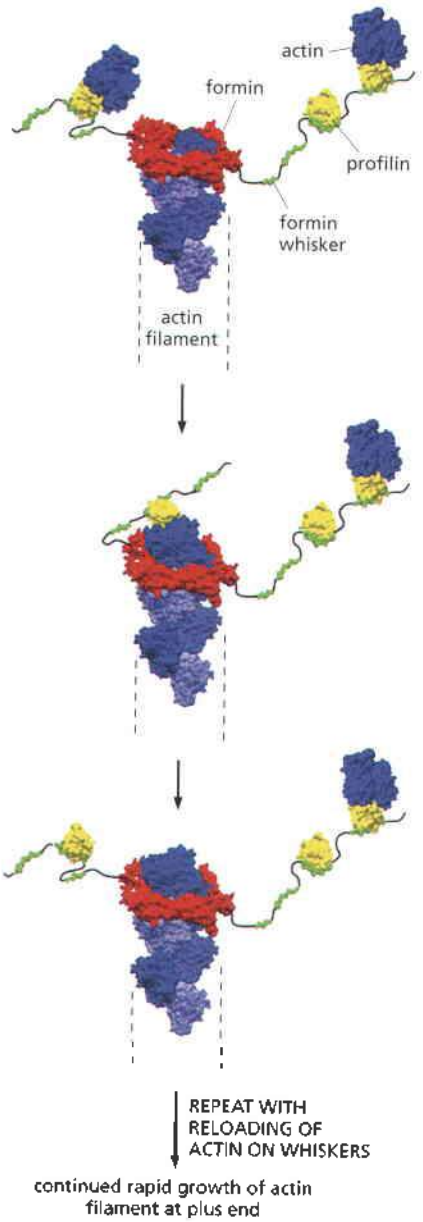


Figure 16–38 Profilin and formins. Some members of the formin protein family have unstructured domains or "whiskers" that contain several binding sites for profilin or the profilin-actin complex. These flexible domains serve as a staging area for addition of actin to the growing plus end of the actin filament when formin is bound. Under some conditions, this can enhance the rate of actin filament elongation so that filament growth is faster than that expected for a diffusion-controlled reaction, and faster in the presence of formin and profilin than the rate for pure actin alone (see also Figure 3–80C).

motile structures such as filopodia and lamellipodia (see below). Besides binding to actin and phospholipids, profilin also binds to various other intracellular proteins that have domains rich in proline; these proteins can also help to localize profilin to sites that require rapid actin assembly.

As it does with actin monomers, the cell sequesters unpolymerized tubulin subunits to maintain the subunit pool at a level substantially higher than the critical concentration. One molecule of the small protein *stathmin* binds to two tubulin heterodimers and prevents their addition onto the ends of microtubules. Stathmin thus decreases the effective concentration of tubulin subunits that are available for polymerization (an action analogous to that of the drug colchicine). Furthermore, stathmin enhances the likelihood that a growing microtubule will undergo the catastrophic transition to the shrinking state. Phosphorylation of stathmin inhibits its binding to tubulin, and signals that cause stathmin phosphorylation can increase the rate of microtubule elongation and suppress dynamic instability. Cancer cells frequently overexpress stathmin, and the increased rate of microtubule turnover that results is thought to contribute to the characteristic change in cell shape associated with malignant transformation.

Severing Proteins Regulate the Length and Kinetic Behavior of Actin Filaments and Microtubules

In some situations, a cell may break an existing long filament into many smaller filaments. This generates a large number of new filament ends: one long filament with just one plus end and one minus end might be broken into dozens of short filaments, each with its own minus end and plus end. Under some intracellular conditions, these newly formed ends nucleate filament elongation, and in this case severing accelerates the assembly of new filament structures. Under other conditions, severing promotes the depolymerization of old filaments, speeding up the depolymerization rate by tenfold or more. In addition, severing filaments changes the physical and mechanical properties of the cytoplasm: stiff, large bundles and gels become more fluid when the filaments are severed.

To sever a microtubule, thirteen longitudinal bonds must be broken, one for each protofilament. The protein *katanin*, named after the Japanese word for "sword," accomplishes this demanding task (**Figure 16–39**). Katanin is made up of two subunits, a smaller subunit that hydrolyzes ATP and performs the actual severing, and a larger one that directs katanin to the centrosome. Katanin releases microtubules from their attachment to a microtubule organizing center, and it is thought to have an important role in the rapid microtubule depolymerization observed at the poles of spindles during meiosis and mitosis. It may also be involved in microtubule release and depolymerization in proliferating cells in interphase and in postmitotic cells such as neurons.

In contrast to microtubule severing by katanin, which requires ATP, the severing of actin filaments does not require an extra energy input. Most actin-severing proteins are members of the *gelsolin superfamily*, whose severing activity is activated by high levels of cytosolic Ca^{2+}. Gelsolin has subdomains that bind to two different sites on the actin subunit, one exposed on the surface of the filament and one that is normally hidden in the longitudinal bond to the next subunit in the protofilament. According to one model for gelsolin severing, gelsolin binds on the side of an actin filament and waits until a thermal fluctuation happens to create a small gap between neighboring subunits in the protofilament; gelsolin then insinuates its subdomain into the gap, breaking the filament.

Proteins That Bind Along the Sides of Filaments Can Either Stabilize or Destabilize Them

Once a cytoskeletal filament is formed by nucleation and elongated from the subunit pool, a set of proteins that bind along the sides of the polymer may alter the filament's stability and mechanical properties. Different filament-associated proteins use their binding energy to either lower or raise the free energy of the polymer state, and they thereby either stabilize or destabilize the polymer, respectively.

Proteins that bind along the sides of microtubules are collectively called **microtubule-associated proteins**, or **MAPs**. Like the drug taxol, MAPs can stabilize microtubules against disassembly. A subset of MAPs can also mediate the interaction of microtubules with other cell components. This subset is prominent in neurons, where stabilized microtubule bundles form the core of the axons and dendrites that extend from the cell body (**Figure 16–40**). These MAPs have at least one domain that binds to the microtubule surface and another that projects outward. The length of the projecting domain can determine how closely MAP-coated microtubules pack together, as demonstrated in cells engineered to overproduce different MAPs. Cells overexpressing *MAP2*, which has a long projecting domain, form bundles of stable microtubules that are kept widely spaced, while cells overexpressing *tau*, a MAP with a much shorter projecting domain, form bundles of more closely packed microtubules (**Figure 16–41**). Tau binding to filaments can also regulate the transport of membrane-enclosed organelles driven by molecular motors, which we will discuss later.

MAPs are the targets of several protein kinases, and the resulting phosphorylation of a MAP can have a primary role in controlling both its activity and localization inside cells. Among the important protein kinases that can regulate MAPs are those that are turned on and off as cells progress through the cell cycle (discussed in Chapter 17). In particular, MAP activities regulate the changes in microtubule dynamics that occur as the cell rearranges its microtubule cytoskeleton to form the mitotic spindle in preparation for chromosome segregation (see Figure 16–2).

In addition to binding along the sides of microtubules, tau protein forms its own helical filaments when present at sufficiently high concentrations. The nerve cell cytoplasm in the brains of people with Alzheimer's disease contains large aggregates of tau filaments, called neurofibrillary tangles. It is not yet clear whether these tangles of tau are a cause or a consequence of the neurodegeneration associated with this disease.

The binding of accessory proteins along their sides also affects actin filaments. Selected actin filaments in most cells are stabilized by the binding of *tropomyosin*, an elongated protein that binds simultaneously to seven adjacent actin subunits in one protofilament. The binding of tropomyosin along an actin filament can prevent the filament from interacting with other proteins; for this reason, the regulation of tropomyosin binding is an important step in muscle contraction, as we discuss later (see Figure 16–78).

Another important actin-filament binding protein present in all eucaryotic cells is *cofilin*, which destabilizes actin filaments. Also called *actin depolymerizing factor*, cofilin is unusual in that it binds to actin in both the filament and free subunit forms. Cofilin binds along the length of the actin filament, forcing the filament to twist a little more tightly (**Figure 16–42**). This mechanical stress

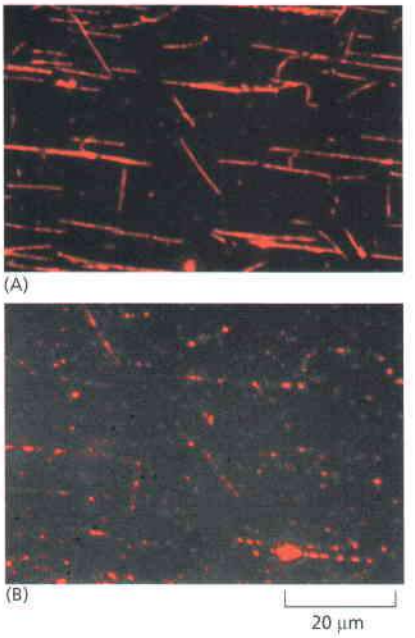


Figure 16–39 Microtubule severing by katanin. Taxol-stabilized, rhodamine-labeled microtubules were adsorbed on the surface of a glass slide, and purified katanin was added along with ATP. (A) There are a few breaks in the microtubules 30 seconds after the addition of katanin. (B) The same field 3 minutes after the addition of katanin. The filaments have been severed in many places, leaving a series of small fragments at the previous locations of the long microtubules. (From J.J. Hartman et al., *Cell* 93:277–287, 1998. With permission from Elsevier.)

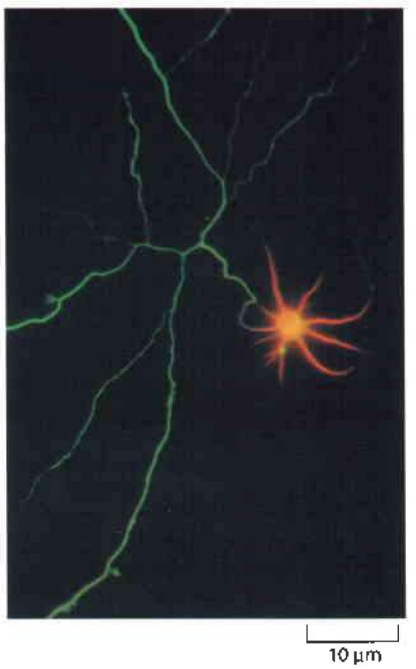


Figure 16–40 Localization of MAPs in axon and dendrites of a neuron. This immunofluorescence micrograph shows the distribution of tau staining *(green)* and MAP2 staining *(orange)* in a hippocampal neuron in culture. Whereas tau staining is confined to the axon (long and branched in this neuron), MAP2 staining is confined to the cell body and its dendrites. The antibody used here to detect tau binds only to unphosphorylated tau; phosphorylated tau is also present in dendrites. (Courtesy of James W. Mandell and Gary A. Banker.)

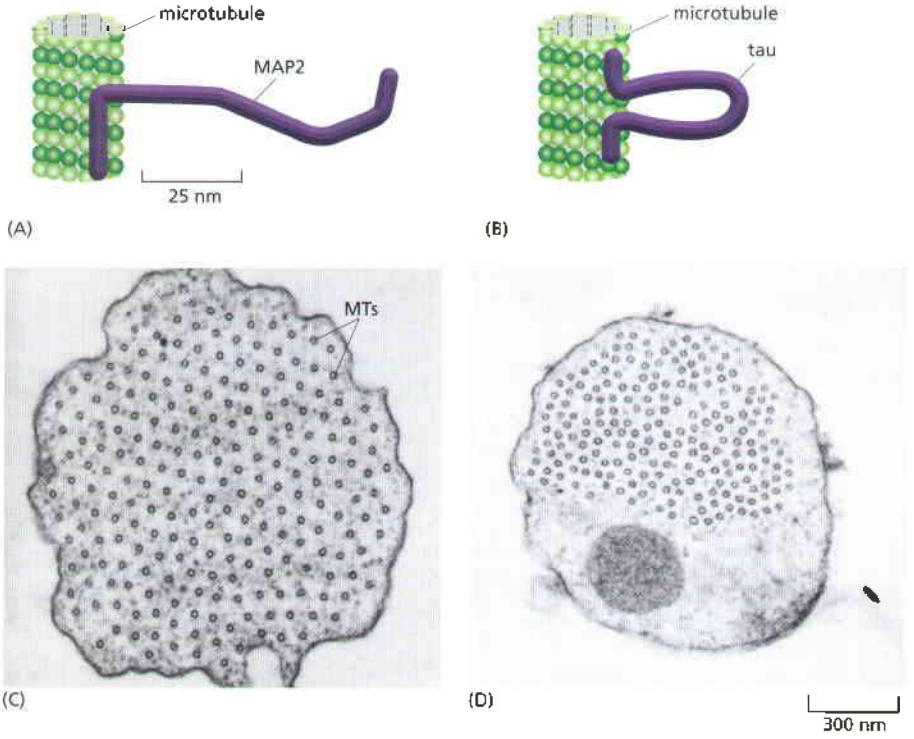


Figure 16–41 Organization of microtubule bundles by MAPs. (A) MAP2 binds along the microtubule lattice at one of its ends and extends a long projecting arm with a second microtubule-binding domain at the other end. (B) Tau binds to the microtubule lattice at both its N- and C-termini, with a short projecting loop. (C) Electron micrograph showing a cross section through a microtubule bundle in a cell overexpressing MAP2. The regular spacing of the microtubules (MTs) in this bundle result from the constant length of the projecting arms of the MAP2. (D) Similar cross section through a microtubule bundle in a cell overexpressing tau. Here the microtubules are spaced more closely together than they are in (C) because of tau's relatively short projecting arm. (C and D, courtesy of V. Chen et al., *Nature* 360:674–647, 1992. With permission from Macmillan Publishers Ltd.)

weakens the contacts between actin subunits in the filament, making the filament brittle and more easily severed by thermal motions. In addition, it makes it much easier for an ADP-actin subunit to dissociate from the minus end of the filament. These activities greatly accelerate actin filament disassembly. As a result, most of the actin filaments inside cells are much shorter-lived than are filaments formed from pure actin in a test tube. Actin filaments can be protected from cofilin by tropomyosin binding.

Cofilin binds preferentially to ADP-containing actin filaments rather than to ATP-containing filaments. Since ATP hydrolysis is usually slower than filament assembly, the newest actin filaments in the cell still contain mostly ATP and are resistant to depolymerization by cofilin. Cofilin therefore efficiently dismantles the older filaments in the cell, ensuring that all actin filaments turn over rapidly. As we will discuss later, the cofilin-mediated disassembly of old but not new actin filaments is critical for the polarized, directed growth of the actin network responsible for unidirectional cell crawling.

Proteins That Interact with Filament Ends Can Dramatically Change Filament Dynamics

As we have just seen, proteins that bind along the side of a filament can change the filament's dynamic behavior. For maximum effect, however, these proteins often need to coat the filament completely, and this means they have to be present at fairly high stoichiometries (for example, about one tropomyosin for every

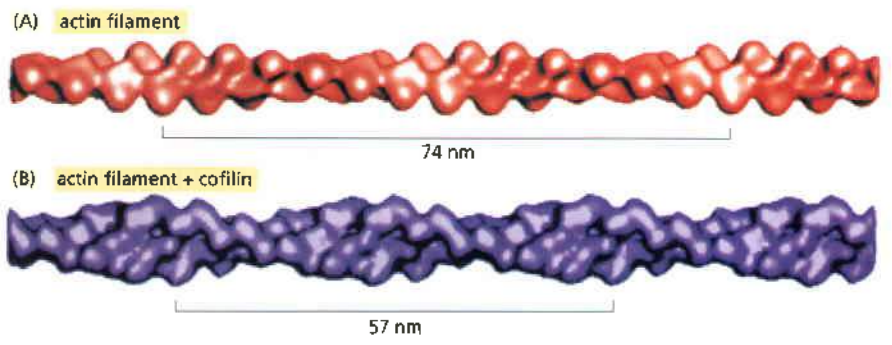


Figure 16–42 Twisting of an actin filament induced by cofilin. (A) Three-dimensional reconstruction from cryo-electron micrographs of filaments made of pure actin. The bracket shows the span of two twists of the actin helix. (B) Reconstruction of an actin filament coated with cofilin, which binds in a 1:1 stoichiometry to actin subunits all along the filament. Cofilin is a small protein (14 kilodaltons) compared to actin (43 kilodaltons), and so the filament appears only slightly thicker. The energy of cofilin binding serves to deform the actin filament lattice, twisting it more tightly and reducing the distance spanned by each twist of the helix. (From A. McGough et al., *J. Cell Biol.* 138:771–781, 1997. With permission from The Rockefeller University Press.)

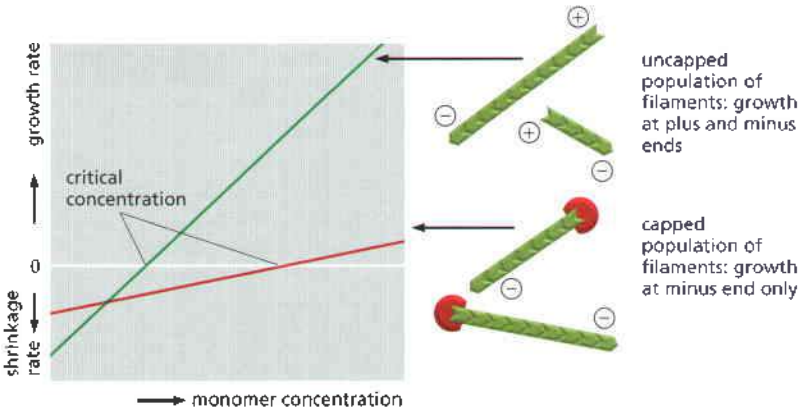


Figure 16–43 Filament capping and its effects on filament dynamics. A population of uncapped filaments adds and loses subunits at both the plus and minus ends, resulting in rapid growth or shrinkage, depending on the concentration of available free monomers (*green* line). In the presence of a protein that caps the plus end (*red* line), only the minus end is able to add or lose subunits; consequently, filament growth will be slower at all monomer concentrations above the critical concentration, and filament shrinkage will be slower at all monomer concentrations below the critical concentration. In addition, the critical concentration for the population shifts to that of the filament minus end.

seven actin subunits, one tau for every four tubulin subunits, or one cofilin for every actin subunit). In contrast, proteins that bind preferentially to the ends of filaments can have dramatic effects on filament dynamics even when they are present at very low levels. Since subunit addition and loss occur primarily at filament ends, one molecule of such a protein per actin filament (typically one per about 200–500 actin monomers) can be enough to transform the architecture of an actin filament network.

As previously discussed, an actin filament that ceases elongation and is not specifically stabilized by the cell can depolymerize rapidly: it can lose subunits from either its plus or its minus end, once the actin molecules at that end have hydrolyzed their ATP to convert to the D form. The most rapid changes, however, occur at the plus end. The binding of a plus end *capping protein* stabilizes an actin filament at its plus end, which greatly slows the rates of both filament growth and filament depolymerization by making the plus end inactive (**Figure 16–43**). Indeed, most of the actin filaments in a cell are capped at their plus end by proteins such as *CapZ* (named for its location in the muscle Z band, see below; it is also called Capping Protein). At the minus end, an actin filament may be capped by remaining bound to the ARP complex that was responsible for its nucleation, although it is possible that many of the actin filament minus ends in typical cells are released from the ARP complex and are uncapped.

In muscle cells, where actin filaments are exceptionally long-lived, the filaments are specially capped at both ends—by CapZ at the plus end and by *tropomodulin* at the minus end. Tropomodulin binds only to the minus end of actin filaments that have been coated by tropomyosin and have thereby already been somewhat stabilized.

Different Kinds of Proteins Alter the Properties of Rapidly Growing Microtubule Ends

The end of a microtubule, with thirteen protofilaments in a hollow ring (see Figure 16–11), is a much larger and more complex structure than the end of an actin filament, with many more possibilities for accessory protein action. We have already discussed an important microtubule capper: the γ-tubulin ring complex (γ-TuRC), which both nucleates the growth of microtubules at an organizing center and caps their minus ends. Another true capping protein for microtubules is the special protein complex found at the ends of the microtubules in cilia (discussed later), where microtubules are both stable and uniform in length.

Some proteins that act at the ends of microtubules have crucial roles beyond those expected for a simple capping protein. In particular, they can have dramatic effects on the dynamic instability of microtubules (see Figure 16–16). They can influence the rate at which a microtubule switches from a growing to a shrinking state (the frequency of catastrophes) or from a shrinking to a growing state (the frequency of rescues). For example, a family of kinesin-related proteins known as *catastrophe factors* significantly increases the catastrophe rate

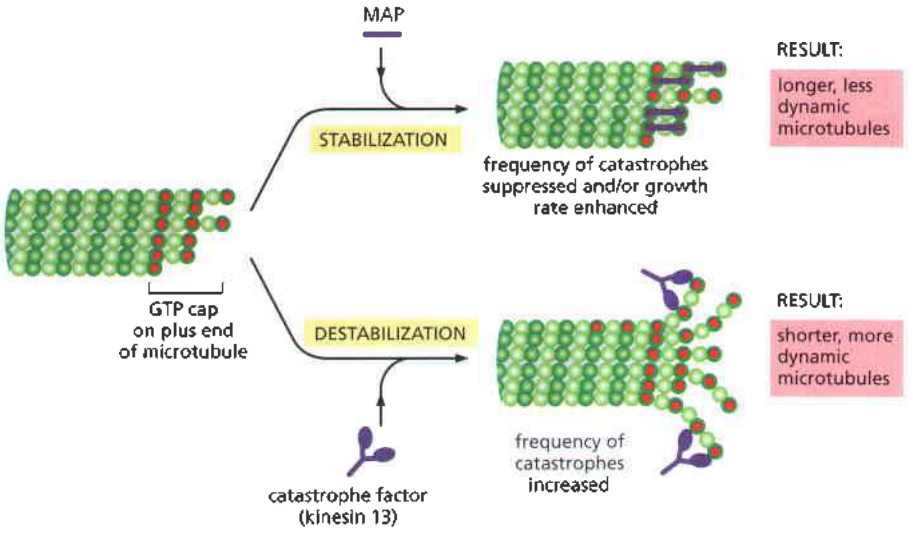
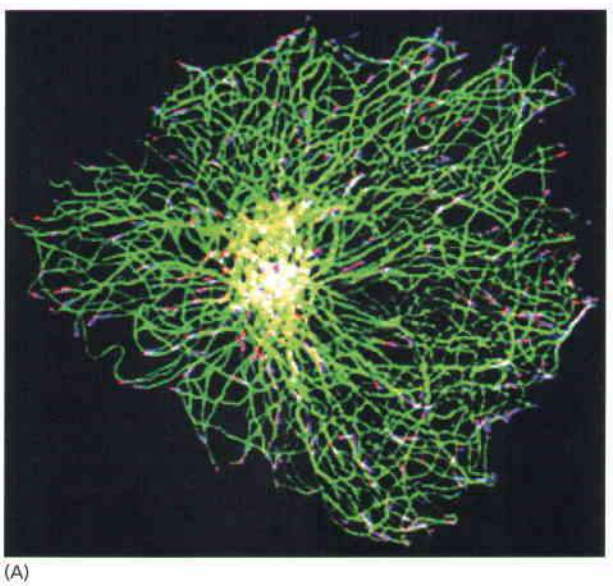


Figure 16–44 The effects of proteins that bind to microtubule ends. The transition between microtubule growth and microtubule shrinking is controlled in cells by special proteins. A MAP such as XMAP215 stabilizes the end of a growing microtubule by its preferential binding there. Opposing its action are catastrophe factors such as kinesin-13, a member of the kinesin motor protein superfamily (discussed later).

(these proteins are members of the kinesin-13 family; see Figure 16–58). They bind specifically to microtubule ends and seem to pry protofilaments apart, lowering the normal activation energy barrier that prevents a microtubule from springing apart into the curved protofilament characteristic of the shrinking state (see Figure 16–16C). Opposing their actions are MAPs such as the ubiquitous *XMAP215* that has close homologs in organisms that range from yeast to humans (XMAP stands for *Xenopus* microtubule-associated protein, and the number refers to its molecular mass in kilodaltons). This protein has a special ability to stabilize free microtubule ends and inhibit their switch from a growing to a shrinking state. The phosphorylation of XMAP215 during mitosis inhibits its activity and shifts the balance of its competition with catastrophe factors (**Figure 16–44**). The shift results in a tenfold increase in the dynamic instability of microtubules observed during mitosis, a transition that is critical for the efficient construction of the mitotic spindle (see Figure 17–33).

In many cells, the minus ends of microtubules are stabilized by association with the centrosome, or else serve as microtubule depolymerization sites. The plus ends, in contrast, efficiently explore and probe the entire cell space. Microtubule-associated proteins called *plus-end tracking proteins (+TIPs)* accumulate at these active ends, and appear to rocket around the cell as passengers at the ends of rapidly growing microtubules, dissociating from the ends when the microtubules begin to shrink (**Figure 16–45**).


(A)

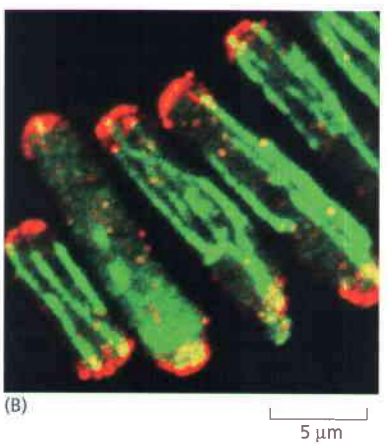


Figure 16–45 +TIP proteins found at the growing plus ends of microtubules. <TAAT> (A) In an epithelial cell grown in tissue culture, each microtubule (*green*) has a growing plus end which is associated with the +TIP protein EB1 (*red*). (B) In the rod-shaped fission yeast *Schizosaccharomyces pombe*, the plus ends of the microtubules (*green*) are associated with the homolog of EB1 (*red*) at the two poles of the rod-shaped cells. (A, from A. Akhmanova and C.C. Hoogenraad, *Curr. Opin. Cell Biol.* 17:47–54, 2005. With permission from Elsevier; B, courtesy of Ken Sawin.)

Some of the +TIPs, such as the kinesin-related catastrophe factors and XMAP215 mentioned above, modulate the growth and shrinkage of the microtubule end to which they are attached. Others control microtubule positioning by helping to capture and stabilize the growing microtubule end at the location of specific target proteins in the cell cortex. EB1, a +TIP present in both yeasts and humans, for example, is essential for yeast mitotic spindle positioning, directing the growing plus ends of yeast spindle microtubules to a specific docking region in the yeast bud and then helping to anchor them there.

Filaments Are Organized into Higher-Order Structures in Cells

So far, we have described how cells use accessory proteins to regulate the location and dynamic behavior of cytoskeletal filaments. These proteins can nucleate filament assembly, bind to the ends or sides of the filaments, or bind to the free subunits of filaments. But in order for the cytoskeletal filaments to form a useful intracellular scaffold that gives the cell mechanical integrity and determines its shape, the individual filaments must be organized and attached to one another in larger-scale structures. The centrosome is one example of such a cytoskeletal organizer; in addition to nucleating the growth of microtubules, it holds them together in a defined geometry, with all of the minus ends buried in the centrosome and the plus ends pointing outward. In this way, the centrosome creates the astral array of microtubules that is able to find the center of each cell (see Figure 16–32).

Another mechanism that cells use to organize filaments into large structures is filament cross-linking. As described earlier, some MAPs can bundle microtubules together: they have two domains—one that binds along the microtubule side (and thereby stabilizes the filament) and another that projects outward to contact other MAP-coated microtubules. In the actin cytoskeleton, the stabilizing and cross-linking functions are separated. Tropomyosin binds along the sides of actin filaments, but it does not have an outward projecting domain. As we shall see shortly, filament cross-linking is instead mediated by a second group of actin-binding proteins that have only this function. Intermediate filaments are different yet again; they are organized both by a lateral self-association of the filaments themselves and by the cross-linking activity of accessory proteins, as we describe next.

Intermediate Filaments Are Cross-Linked and Bundled Into Strong Arrays

Each individual intermediate filament forms as a long bundle of tetrameric subunits (see Figure 16–19). Many intermediate filaments further bundle themselves by self-association; for example, the neurofilament proteins NF-M and NF-H (see Table 16–1, p. 985) contain a C-terminal domain that extends outward from the surface of the assembled intermediate filament and binds to a neighboring filament. Thus groups of neurofilaments form robust parallel arrays that are held together by multiple lateral contacts, giving strength and stability to the long cell processes of neurons (see Figure 16–22).

Other types of intermediate filament bundles are held together by accessory proteins, such as *filaggrin*, which bundles keratin filaments in differentiating cells of the epidermis to give the outermost layers of the skin their special toughness. *Plectin* is a particularly interesting cross-linking protein. Besides bundling intermediate filaments, it also links the intermediate filaments to microtubules, actin filament bundles, and filaments of the motor protein myosin II (discussed below), as well as helping to attach intermediate filament bundles to adhesive structures at the plasma membrane (**Figure 16–46**).

Mutations in the gene for plectin cause a devastating human disease that combines epidermolysis bullosa (caused by disruption of skin keratin filaments), muscular dystrophy (caused by disruption of desmin filaments), and neurodegeneration (caused by disruption of neurofilaments). Mice lacking a

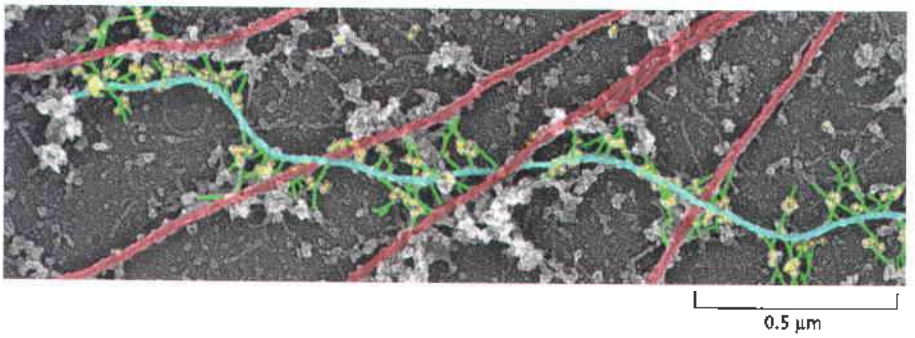


Figure 16–46 Plectin cross-linking of diverse cytoskeletal elements. Plectin *(green)* is seen here making cross-links from intermediate filaments *(blue)* to microtubules *(red)*. In this electron micrograph, the dots *(yellow)* are gold particles linked to anti-plectin antibodies. The entire actin filament network was removed to reveal these proteins. (From T.M. Svitkina and G.G.Borisy, *J. Cell Biol.* 135:991–1007, 1996. With permission from The Rockefeller University Press.)

functional plectin gene die within a few days of birth, with blistered skin and abnormal skeletal and heart muscles. Thus, although plectin may not be necessary for the initial formation and assembly of intermediate filaments, its cross-linking action is required to provide cells with the strength they need to withstand the mechanical stresses inherent to vertebrate life.

Cross-linking Proteins with Distinct Properties Organize Different Assemblies of Actin Filaments

Actin filaments in animal cells are organized into two types of arrays: bundles and weblike (gel-like) networks (**Figure 16–47**). As described earlier, these different structures are initiated by the action of distinct nucleating proteins: the long straight filaments produced by formins make bundles and the ARP complex makes webs. The actin filament cross-linking proteins that help to stabilize and maintain these distinct structures are divided into two classes: *bundling proteins* and *gel-forming proteins*. Bundling proteins cross-link actin filaments into a parallel array, while gel-forming proteins hold two actin filaments together at a large angle to each other, thereby creating a looser meshwork. Both types of cross-linking protein generally have two similar actin-filament-binding sites, which can either be part of a single polypeptide chain or contributed by each of two polypeptide chains held together in a dimer (**Figure 16–48**). The spacing and arrangement of these two filament-binding domains determines the type of actin structure that a given cross-linking protein forms.

Each type of bundling protein also determines which other molecules can interact with an actin filament. Myosin II (discussed later) is the motor protein in stress fibers and other contractile arrays that enables them to contract. The

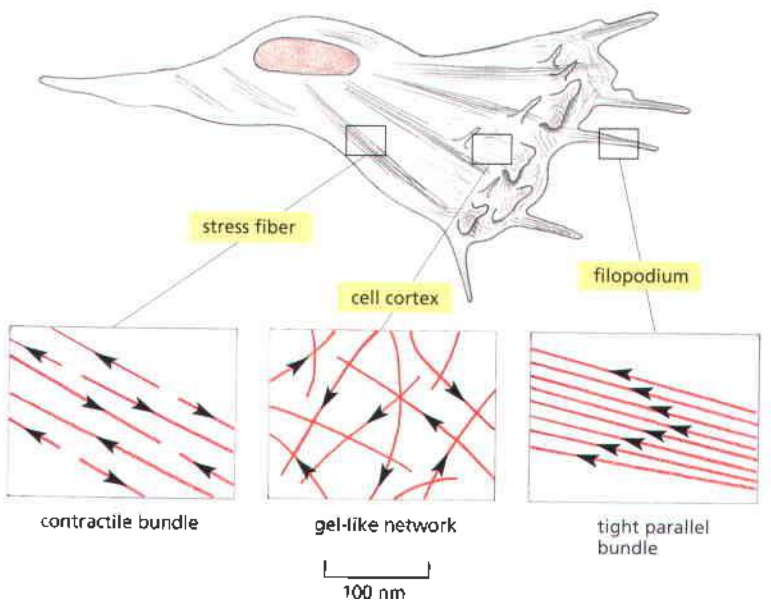


Figure 16–47 Actin arrays in a cell. A fibroblast crawling in a tissue culture dish is shown with three areas enlarged to show the arrangement of actin filaments. The actin filaments are shown in *red*, with arrowheads pointing toward the minus end. Stress fibers are contractile and exert tension. Filopodia are spike-like projections of the plasma membrane that allow a cell to explore its environment. The cortex underlies the plasma membrane.

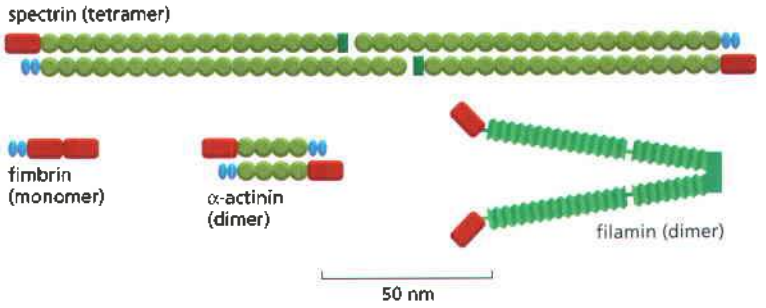
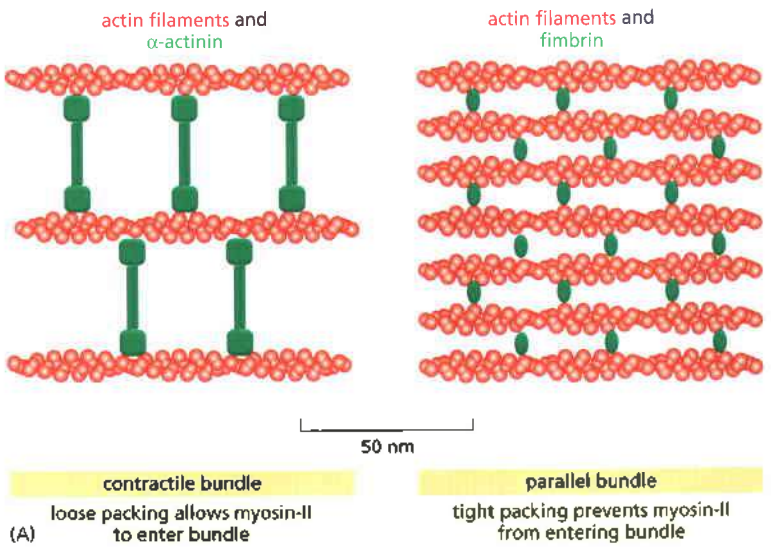


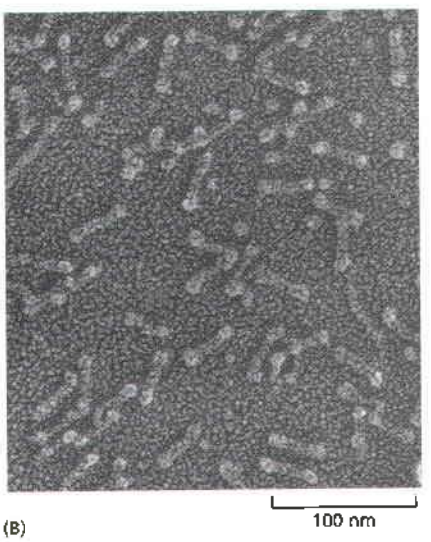
Figure 16–48 The modular structures of four actin-cross-linking proteins. Each of the proteins shown has two actin-binding sites *(red)* that are related in sequence. Fimbrin has two directly adjacent actin-binding sites, so that it holds its two actin filaments very close together (14 nm apart), aligned with the same polarity (see Figure 16–49A). The two actin-binding sites in α-actinin are separated by a spacer around 30 nm long, so that it forms more loosely packed actin bundles (see Figure 16–49A). Filamin has two actin-binding sites with a V-shaped linkage between them, so that it cross-links actin filaments into a network with the filaments oriented almost at right angles to one another (see Figure 16–51). Spectrin is a tetramer of two α and two β subunits, and the tetramer has two actin-binding sites spaced about 200-nm apart (see Figure 10–41).

very close packing of actin filaments caused by the small monomeric bundling protein *fimbrin* apparently excludes myosin, and thus the parallel actin filaments held together by fimbrin are not contractile; on the other hand, the looser packing caused by the larger dimeric bundling protein *α-actinin* allows myosin molecules to enter, making stress fibers contractile (**Figure 16–49**). Because of the very different spacing between the actin filaments, bundling by fimbrin automatically discourages bundling by α-actinin, and vice-versa, so that the two types of bundling protein are themselves mutually exclusive.

Villin is another bundling protein that, like fimbrin, has two actin-filament-binding sites very close together in a single polypeptide chain. Villin (together with fimbrin) helps cross-link the 20 to 30 tightly bundled actin filaments found in microvilli, the finger-like extensions of the plasma membrane on the surface of many epithelial cells (**Figure 16–50**). A single absorptive epithelial cell in the human small intestine, for example, has several thousand microvilli on its apical surface. Each is about 0.08 μm wide and 1 μm long, making the cell's absorptive surface area about 20 times greater than it would be without microvilli. When villin is introduced into cultured fibroblasts, which do not normally contain villin and have only a few small microvilli, the existing microvilli become greatly elongated and stabilized, and new ones are induced. The actin filament core of the microvillus is attached to the plasma membrane along its sides by lateral sidearms made of *myosin I* (discussed later), which has a binding site for filamentous actin on one end and a domain that binds lipids on the other end. These two types of cross-linkers, one binding actin filaments to each other and the other binding these filaments to the membrane, seem to be sufficient to form microvilli on cells. Interestingly, when the gene for villin is disrupted in a mouse, the intestinal microvilli form with apparently normal morphology, indicating that other bundling proteins provide sufficient redundant function for this purpose. However, the remodeling of intestinal microvilli in response to certain kinds of stress or starvation is impaired.

Figure 16–49 The formation of two types of actin filament bundles. (A) α-actinin, which is a homodimer, cross-links actin filaments into loose bundles, which allow the motor protein myosin II (not shown) to participate in the assembly. Fimbrin cross-links actin filaments into tight bundles, which exclude myosin. Fimbrin and α-actinin tend to exclude one another because of the very different spacing of the actin filament bundles that they form. (B) Electron micrograph of purified α-actinin molecules. (B, courtesy of John Heuser.)





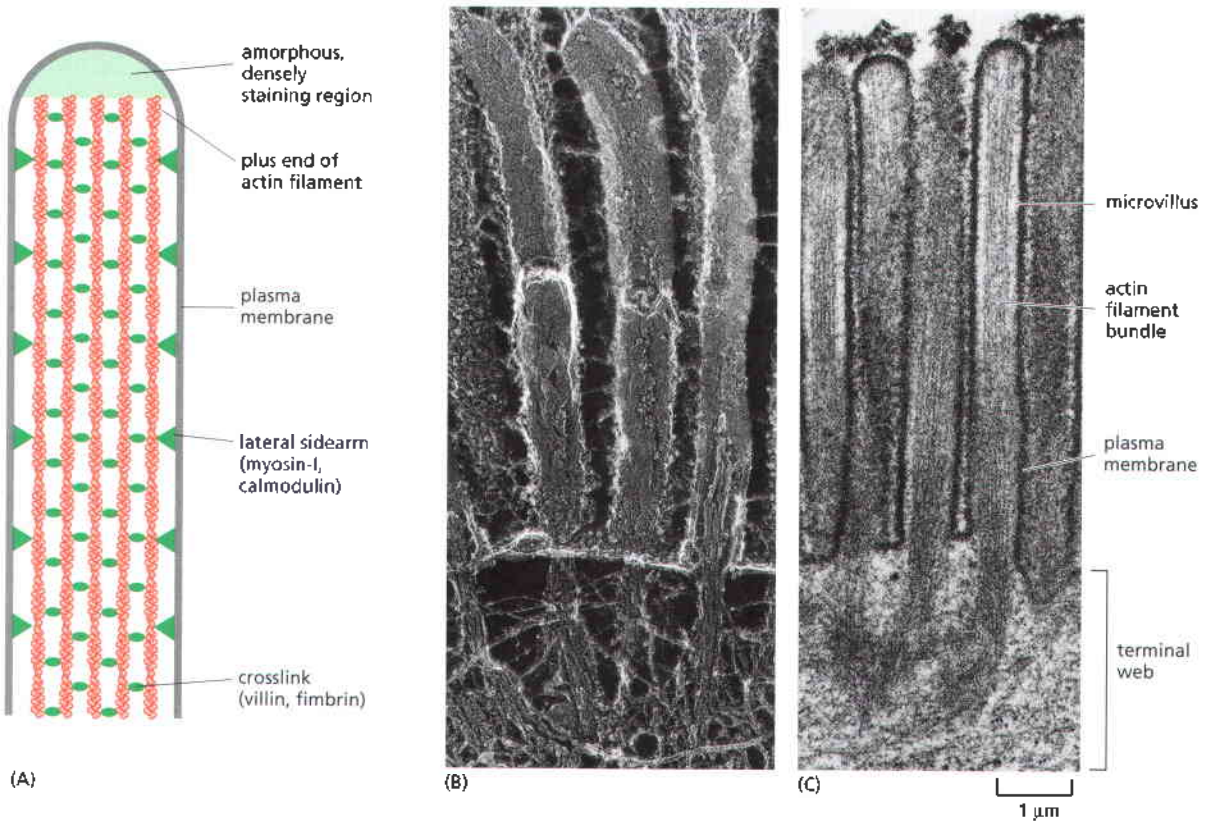


Figure 16–50 **A microvillus.** (A) A bundle of parallel actin filaments cross-linked by the actin-bundling proteins villin and fimbrin forms the core of a microvillus. Lateral sidearms (composed of myosin I and the Ca^{2+}-binding protein calmodulin) connect the sides of the actin filament bundle to the overlying plasma membrane. All the plus ends of the actin filaments are at the tip of the microvillus, where they are embedded in an amorphous, densely staining substance of unknown composition. (B) Freeze-fracture electron micrograph of the apical surface of an intestinal epithelial cell, showing microvilli. Actin bundles from the microvilli extend down into the cell and are rooted in the terminal web, where they are linked together by a complex set of proteins that includes spectrin and myosin II. Below the terminal web is a layer of intermediate filaments. (C) Thin section electron micrograph of microvilli. (B, courtesy of John Heuser; C, from P.T. Matsudaira and D.R. Burgess, *Cold Spring Harb. Symp. Quant. Biol.* 46:845–854, 1985. With permission from Cold Spring Harbor Laboratory Press.)

Filamin and Spectrin Form Actin Filament Webs

The various bundling proteins that we have discussed so far have straight, stiff connections between their two actin-filament-binding domains, and they tend to align filaments in parallel bundles. In contrast, those actin cross-linking proteins that have either a flexible or a stiff, bent connection between their two binding domains form actin filament webs or gels, rather than actin bundles.

Any cross-linking protein that has its two actin-binding domains joined by a long bent linkage can form three-dimensional actin gels. *Filamin* (see Figure 16–48) promotes the formation of a loose and highly viscous gel by clamping together two actin filaments roughly at right angles (**Figure 16–51**). Cells require the actin gels formed by filamin in order to extend the thin sheet-like membrane projections called *lamellipodia* that help them to crawl across solid surfaces. Filamin is lacking in some types of cancer cells, especially some malignant melanomas (pigment-cell cancers). These cells cannot crawl properly, and instead they protrude disorganized membrane blebs (**Figure 16–52**). Losing filamin is bad for the melanoma cells but good for the melanoma patient; because of the cells' inability to crawl, melanoma cells that have lost filamin expression are less invasive than similar melanoma cells that still express filamin, and, as a result, the cancer is much less likely to metastasize.

A very different well-studied web-forming protein is *spectrin*, which was first identified in red blood cells. Spectrin is a long, flexible protein made out of four elongated polypeptide chains (two α subunits and two β subunits), arranged so that the two actin-filament-binding sites are about 200 nm apart (compared

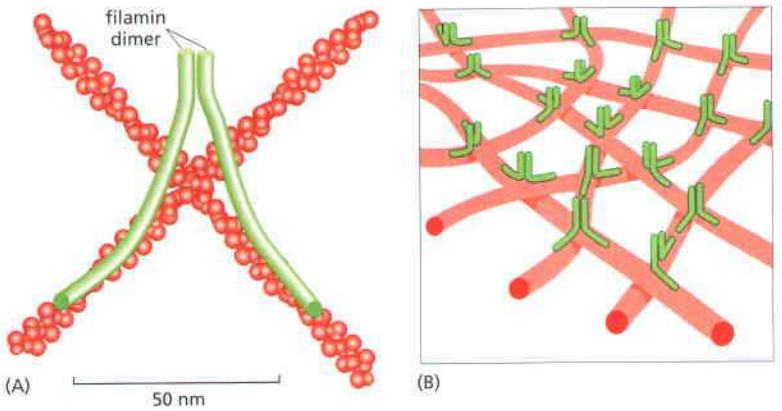


Figure 16–51 Filamin cross-links actin filaments into a three-dimensional network with the physical properties of a gel. (A) Each filamin homodimer is about 160 nm long when fully extended and forms a flexible, high-angle link between two adjacent actin filaments. (B) A set of actin filaments cross-linked by filamin forms a mechanically strong web or gel.

with 14 nm for fimbrin and about 30 nm for α-actinin, see Figure 16–48). In the red blood cell, spectrin is concentrated just beneath the plasma membrane, where it forms a two-dimensional web held together by short actin filaments; spectrin links this web to the plasma membrane because it has separate binding sites for peripheral membrane proteins, which are themselves positioned near the lipid bilayer by integral membrane proteins (see Figure 10–41). The resulting network creates a stiff cell cortex that provides mechanical support for the overlying plasma membrane, allowing the red blood cell to spring back to its original shape after squeezing through a capillary. Close relatives of spectrin are found in the cortex of most other vertebrate cell types, where they also help to shape and stiffen the surface membrane.

Cytoskeletal Elements Make Many Attachments to Membranes

Actin cytoskeletal structures both stiffen and change the shape of the plasma membrane. We have already discussed two examples: the spectrin-actin web that underlies the plasma membranes and the villin-actin bundles in microvilli that enlarge the absorptive surface area of epithelial cells. The effectiveness of these structures requires specific attachments between the actin filament network and proteins or lipids of the plasma membrane.

The connections of the cortical actin cytoskeleton to the plasma membrane are only partially understood. A widespread family of closely related intracellular proteins, the *ERM* family (named for its first three members, ezrin, radixin, and moesin), contains members that are required for the maintenance of cell polarity and involved in exocytosis and endocytosis. The C-terminal domain of an ERM protein binds directly to the sides of actin filaments. The N-terminal domain binds to the cytoplasmic face of one or more transmembrane glycoproteins, such as CD44, the receptor for the extracellular matrix component hyaluronan.

The attachments between actin and the plasma membrane mediated by ERM proteins are regulated by both intracellular and extracellular signals. ERM proteins can exist in two conformations, an active extended conformation that oligomerizes and binds to both actin and a transmembrane protein, and an inactive folded conformation, in which the N- and C-termini are held together by an intramolecular interaction. Switching to the active conformation can be

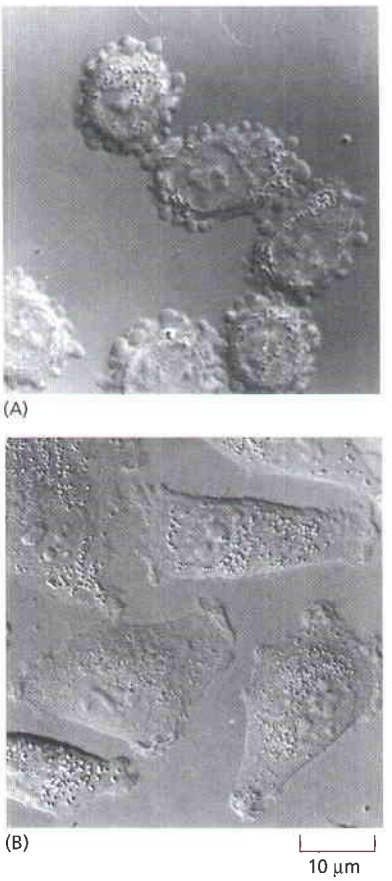


Figure 16–52 Loss of filamin causes abnormal cell motility. (A) A group of melanoma cells that have an abnormally low level of filamin. These cells are not able to make normal lamellipodia and instead are covered with membrane "blebs." As a result, they crawl poorly and tend not to metastasize. (B) The same melanoma cells in which filamin expression has been artificially restored. The cells now make normal lamellipodia and are highly metastatic. This example is one of many demonstrating the profound effect that the presence or absence of a single structural protein can have on cell morphology and motility. (From C. Cunningham et al., *J. Cell Biol.* 136:845–857, 1997. With permission The Rockefeller University Press.)

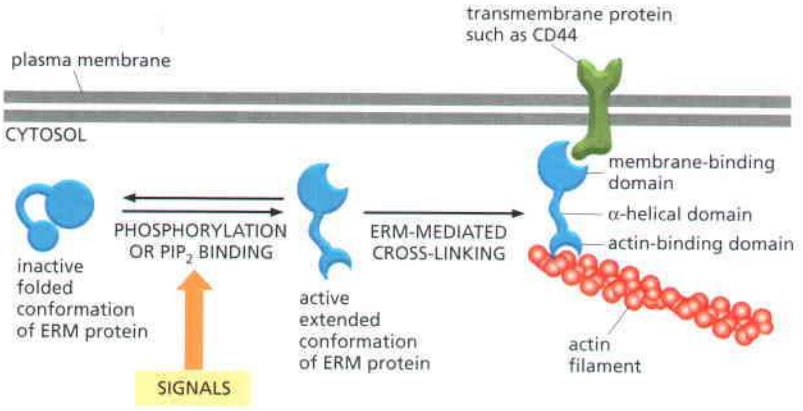


Figure 16–53 **The role of ERM-family proteins in attaching actin filaments to the plasma membrane.** Regulated unfolding of an ERM-family protein, caused by phosphorylation or by binding to PIP_2, exposes two binding sites, one for an actin filament and one for a transmembrane protein. Activation of ERM-family proteins can thereby generate and stabilize cell-surface protrusions that form in response to extracellular signals.

triggered by protein phosphorylation or by binding to PIP_2, either of which can occur, for example, in response to extracellular signals (the ERM proteins are direct targets of several receptor tyrosine kinases). In this way, the properties of the cell cortex become sensitive to a variety of signals received by the cell (Figure 16–53).

The loss of one of the members of the ERM family, called merlin, results in a form of the human genetic disease called *neurofibromatosis*, in which multiple benign tumors develop in the auditory nerves and certain other parts of the nervous system. This is one of many indications of a feedback system that connects cell structural elements to the control of cell growth (see Chapter 17).

The proteins, discussed in this section, that control the assembly and position of actin filaments and microtubules are reviewed in Panel 16–3 (pp. 994–995). Some of these proteins have the additional function of helping to connect the internal structure of a cell to other cells or to an extracellular basement membrane. Both actin filaments and intermediate filaments are critical for these connections, which require the specialized cell-cell junctions and cell-matrix junctions that we will discuss in Chapter 19.

Summary

The varied forms and functions of cytoskeletal filament structures in eucaryotic cells depend on a versatile repertoire of accessory proteins. Each of the three major filament classes (microtubules, intermediate filaments, and actin filaments) has a large dedicated subset of such accessory proteins.

A primary determinant of the sites of cytoskeletal structures is the regulation of the processes that initiate the nucleation of new filaments. In most animal cells, microtubules are nucleated at the centrosome, a complex assembly located near the center of the cell. In contrast, most actin filaments are nucleated near the plasma membrane.

The kinetics of filament assembly and disassembly can be either slowed or accelerated by accessory proteins that bind to either the free subunits or the filaments themselves. Some of these proteins alter filament dynamics by binding to the ends of filaments or by severing the filaments into smaller fragments. Another class of accessory proteins assembles the filaments into larger ordered structures by cross-linking them to one another in geometrically defined ways. Yet other accessory proteins determine the shape and adhesive properties of cells by attaching filaments to the plasma membrane.

MOLECULAR MOTORS

Among the most fascinating proteins that associate with the cytoskeleton are the molecular motors called **motor proteins**. These remarkable proteins bind to a polarized cytoskeletal filament and use the energy derived from repeated cycles of ATP hydrolysis to move steadily along it. Dozens of different motor proteins coexist in every eucaryotic cell. They differ in the type of filament they bind to (either actin or microtubules), the direction in which they move along the filament, and

the "cargo" they carry. Many motor proteins carry membrane-enclosed organelles—such as mitochondria, Golgi stacks, or secretory vesicles—to their appropriate locations in the cell. Other motor proteins cause cytoskeletal filaments to exert tension or to slide against each other, generating the force that drives such phenomena as muscle contraction, ciliary beating, and cell division.

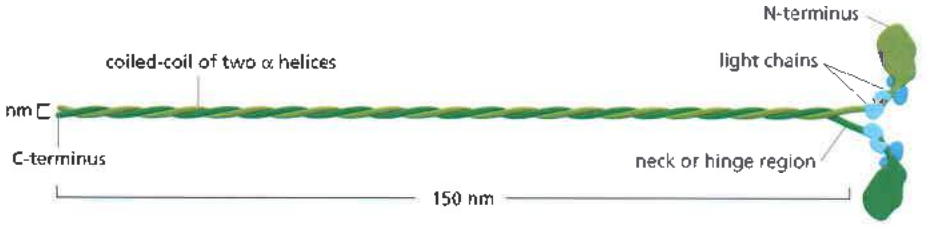
Cytoskeletal motor proteins that move unidirectionally along an oriented polymer track are reminiscent of some other proteins and protein complexes discussed elsewhere in this book, such as DNA and RNA polymerases, helicases, and ribosomes. All of these have the ability to use chemical energy to propel themselves along a linear track, with the direction of sliding dependent on the structural polarity of the track. All of them generate motion by coupling nucleoside triphosphate hydrolysis to a large-scale conformational change in a protein, as explained in Chapter 3 (see Figure 3–77).

The cytoskeletal motor proteins associate with their filament tracks through a "head" region, or *motor domain*, that binds and hydrolyzes ATP. Driven by cycles of nucleotide hydrolysis that produce conformational changes, the proteins cycle between states in which they are bound strongly to their filament tracks and states in which they are unbound. Through a mechanochemical cycle of filament binding, conformational change, filament release, conformational relaxation, and filament rebinding, the motor protein and its associated cargo move one step at a time along the filament (typically a distance of a few nanometers). The motor domain (head) determines the identity of the track and the direction of movement along it, whereas the tail of the motor protein determines the identity of the cargo (and therefore the biological function of the individual motor protein).

In this section, we begin by describing the three groups of cytoskeletal motor proteins: myosins, kinesins, and dyneins. We then describe how they can work to transport membrane-enclosed organelles and mRNAs or to change the shape of structures built from cytoskeletal filaments. In the final section of this chapter, we will examine how a collaboration between motor proteins and the dynamic cytoskeletal filaments described previously generates complex cell behaviors.

Actin-Based Motor Proteins Are Members of the Myosin Superfamily

The first motor protein identified was skeletal muscle **myosin**, which generates the force for muscle contraction. This myosin, called *myosin II* (see below) is an elongated protein that is formed from two heavy chains and two copies of each of two light chains. Each heavy chain has a globular head domain at its N-terminus that contains the force-generating machinery, followed by a very long amino acid sequence that forms an extended coiled-coil that mediates heavy chain dimerization (**Figure 16–54**). The two light chains bind close to the N-terminal



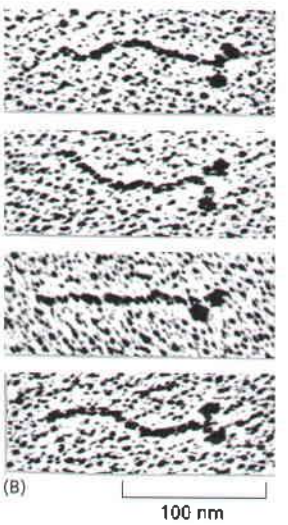


Figure 16–54 **Myosin II.** (A) A myosin II molecule is composed of two heavy chains (each about 2000 amino acids long *(green)* and four light chains *(blue)*. The light chains are of two distinct types, and one copy of each type is present on each myosin head. Dimerization occurs when the two α helices of the heavy chains wrap around each other to form a coiled-coil, driven by the association of regularly spaced hydrophobic amino acids (see Figure 3–9). The coiled-coil arrangement makes an extended rod in solution, and this part of the molecule is called the tail. (B) The two globular heads and the tail can be clearly seen in electron micrographs of myosin molecules shadowed with platinum. (B, courtesy of David Shotton.)

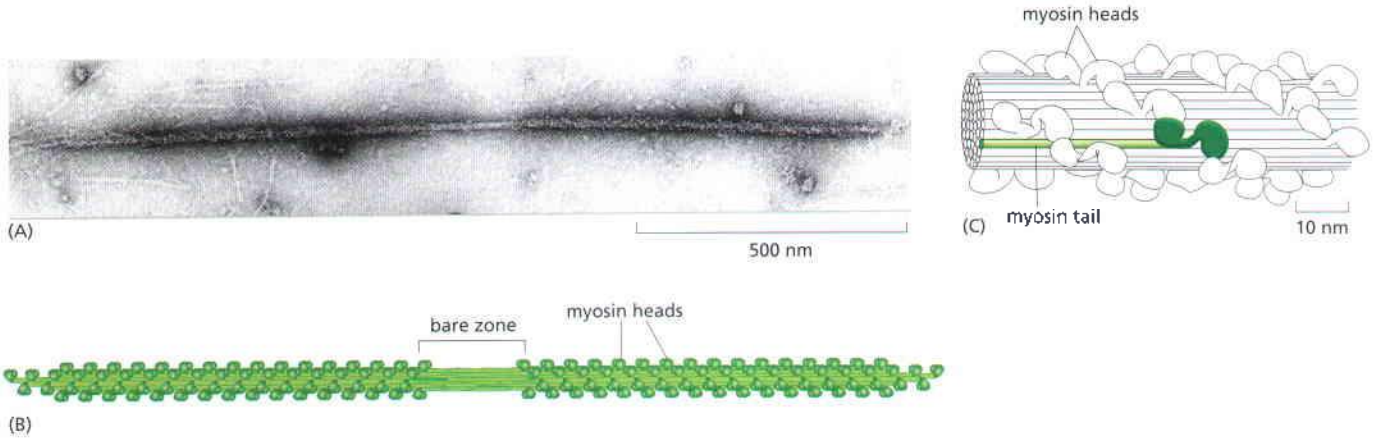


Figure 16–55 The myosin II bipolar thick filament in muscle. (A) Electron micrograph of a myosin II thick filament isolated from frog muscle. Note the central bare zone, which is free of head domains. (B) Schematic diagram, not drawn to scale. The myosin II molecules aggregate by means of their tail regions, with their heads projecting to the outside of the filament. The bare zone in the center of the filament consists entirely of myosin II tails. (C) A small section of a myosin II filament as reconstructed from electron micrographs. An individual myosin molecule is highlighted in *green*. The cytoplasmic myosin II filaments in non-muscle cells are much smaller, although similarly organized (see Figure 16–72). (A, courtesy of Murray Stewart; C, based on R.A. Crowther, R. Padron and R. Craig, *J. Mol. Biol.* 184:429–439, 1985. With permission from Academic Press.)

head domain, while the long coiled-coil tail bundles itself with the tails of other myosin molecules. These tail–tail interactions form large bipolar "thick filaments" that have several hundred myosin heads, oriented in opposite directions at the two ends of the thick filament (**Figure 16–55**).

Each myosin head binds and hydrolyzes ATP, using the energy of ATP hydrolysis to walk toward the plus end of an actin filament. The opposing orientation of the heads in the thick filament makes the filament efficient at sliding pairs of oppositely oriented actin filaments past each other. In skeletal muscle, in which carefully arranged actin filaments are aligned in "thin filament" arrays surrounding the myosin thick filaments, the ATP-driven sliding of actin filaments results in muscle contraction (discussed later). Cardiac and smooth muscle contain myosin II molecules that are similarly arranged, although different genes encode them.

When a muscle myosin is digested by chymotrypsin and papain, the head domain is released as an intact fragment (called S1). The S1 fragment alone can generate filament sliding *in vitro*, proving that the motor activity is contained completely within the head (**Figure 16–56**).

It was initially thought that myosin was present only in muscle, but in the 1970s, researchers found that a similar two-headed myosin protein was also

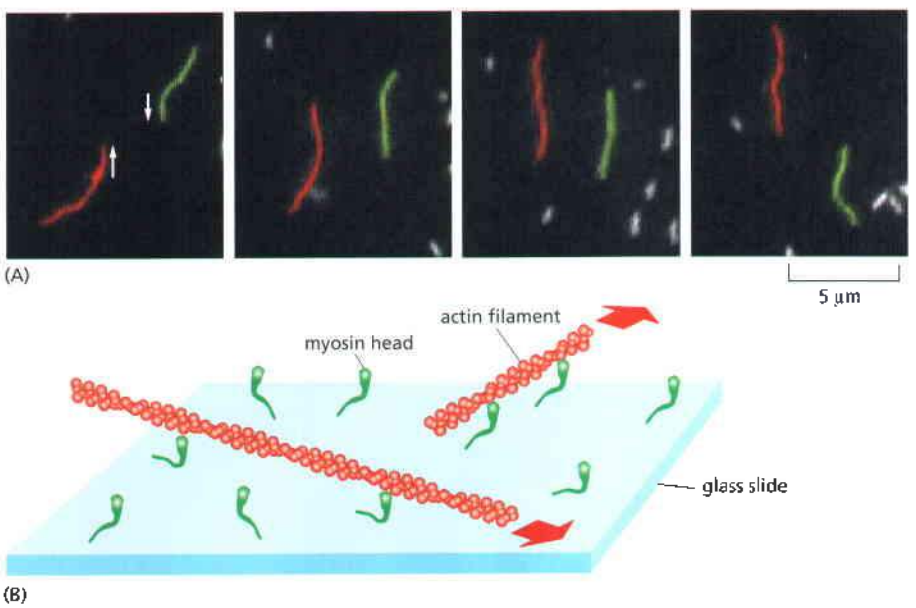


Figure 16–56 Direct evidence for the motor activity of the myosin head. <TTAT> In this experiment, purified S1 myosin heads were attached to a glass slide, and then actin filaments labeled with fluorescent phalloidin were added and allowed to bind to the myosin heads. (A) When ATP was added, the actin filaments began to glide along the surface, owing to the many individual steps taken by each of the dozens of myosin heads bound to each filament. The video frames shown in this sequence were recorded about 0.6 second apart; the two actin filaments shown (one *red* and one *green*) were moving in opposite directions at a rate of about 4 μm/sec. (B) Diagram of the experiment. The large *red* arrows indicate the direction of actin filament movement. (A, courtesy of James Spudich.)

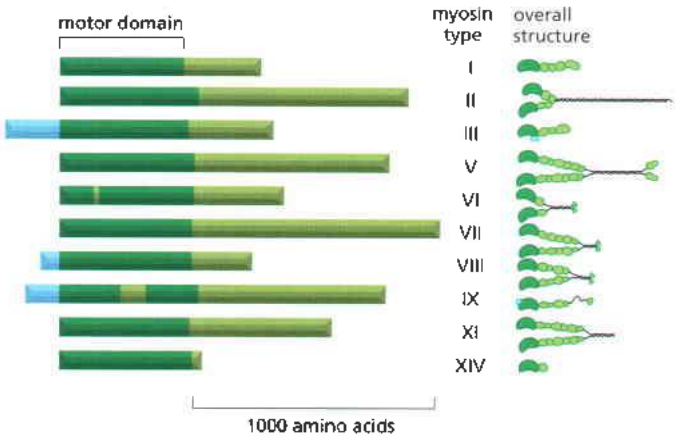


Figure 16–57 Myosin superfamily members. Comparison of the domain structure of the heavy chains of some myosin types. All myosins share similar motor domains (shown in *dark green*), but their C-terminal tails *(light green)* and N-terminal extensions *(light blue)* are very diverse. On the right are depictions of the molecular structure for these family members. Many myosins form dimers, with two motor domains per molecule, but a few (such as I, IX, and XIV) seem to function as monomers, with just one motor domain. Myosin VI, despite its overall structural similarity to other family members, is unique in moving toward the minus end (instead of the plus end) of an actin filament. The small insertion within its motor head domain, not found in other myosins, is probably responsible for this change in direction.

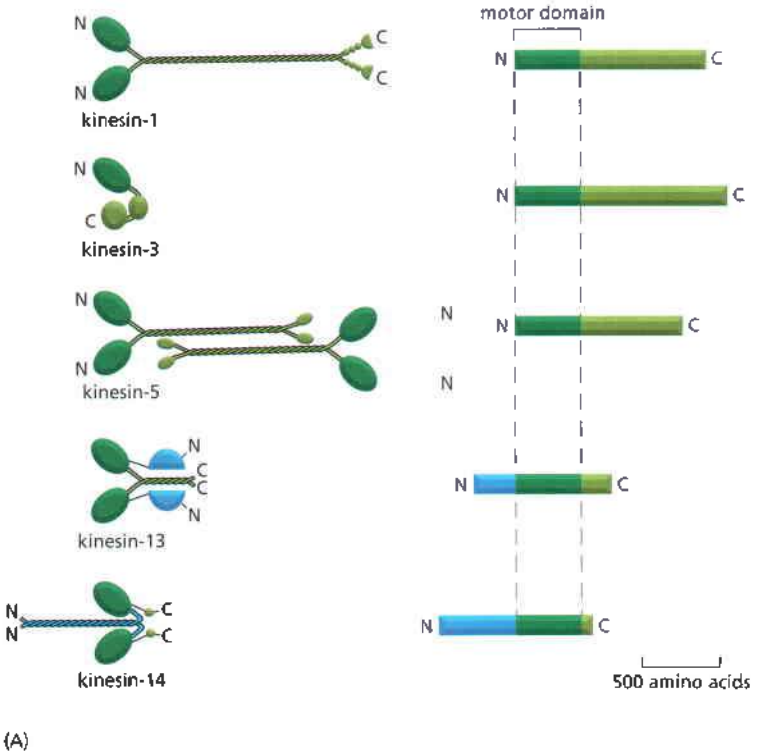
present in nonmuscle cells, including protozoan cells. At about the same time, other researchers found a myosin in the freshwater amoeba *Acanthamoeba castellanii* that was unconventional in having a motor domain similar to the head of muscle myosin but a completely different tail. This molecule seemed to function as a monomer and was named *myosin I* (for one-headed); the conventional myosin was renamed *myosin II* (for two-headed).

Subsequently, many other myosin types were discovered. The heavy chains generally start with a recognizable myosin motor domain at the N-terminus, and then diverge widely with a variety of C-terminal tail domains (**Figure 16–57**). The newly identified types of myosins include a number of one-headed and two-headed varieties that are approximately equally related to myosin I and myosin II, and the nomenclature now reflects their approximate order of discovery (myosin III through at least myosin XVIII). Sequence comparisons among diverse eucaryotes indicate that there are at least 37 distinct myosin families in the superfamily. The myosin tails (and the tails of motor proteins generally) have apparently diversified during evolution to permit the proteins to bind other subunits and to interact with different cargoes.

Some myosins (such as VIII and XI) have been found only in plants, and some have been found only in vertebrates (IX). Most, however, are found in all eucaryotes, suggesting that myosins arose early in eucaryotic evolution. The yeast *Saccharomyces cerevisiae* contains five myosins: two myosin Is, one myosin II, and two myosin Vs. It is tempting to speculate that these three types of myosins are necessary for a eucaryotic cell to survive and that other myosins perform more specialized functions, particularly in multicellular organisms. The nematode *C. elegans*, for example, has at least 15 myosin genes, representing at least seven structural classes; the human genome includes about 40 myosin genes. Nine of the human myosins are expressed primarily or exclusively in the hair cells of the inner ear, and mutations in five of them are known to cause hereditary deafness. These extremely specialized myosins are important for the construction and function of the complex and beautiful bundles of actin-rich stereocilia that are found on the apical surface of these cells (see Figure 9–50); these tilt in response to sound and convert sound waves into electrical signals (discussed in Chapter 23).

All of the myosins except one move toward the plus end of an actin filament, although they do so at different speeds. The exception is myosin VI, which moves toward the minus end.

The exact functions for most of the myosins remain to be determined. Myosin V is involved in vesicle and organelle transport. Myosin II is associated with contractile activity in both muscle and nonmuscle cells. It is generally required for cytokinesis (the pinching apart of a dividing cell into two daughters), as well as for the forward translocation of the body of a cell during cell migration. The myosin I proteins often contain a second actin-binding site or a membrane-binding site in their tails, and they are generally involved in intracellular organization—including the protrusion of actin-rich structures at the cell surface, as discussed earlier for the construction of microvilli (see Figure 16–50).



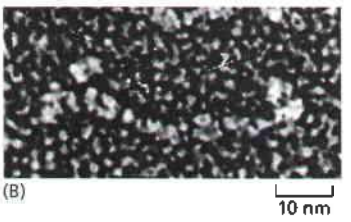


Figure 16–58 Kinesin and kinesin-related proteins. (A) Structures of five kinesin superfamily members. As in the myosin superfamily, only the motor domains are conserved. Kinesin-1 has the motor domain at the N-terminus of the heavy chain. The middle domain forms a long coiled-coil, mediating dimerization. The C-terminal domain forms a tail that attaches to cargo, such as a membrane-enclosed organelle. Kinesin-3 represents an unusual class of kinesins that seem to function as monomers and move membrane-enclosed organelles along microtubules. Kinesin-5 forms tetramers where two dimers associate by their tails. The bipolar kinesin-5 tetramer is able to slide two microtubules past each other, analogous to the activity of the bipolar thick filaments formed by myosin II. Kinesin-13 has its motor domain located in the middle of the heavy chain. It is a member of a family of kinesins that have lost typical motor activity and instead bind to microtubule ends to increase dynamic instability of microtubules; they are therefore called catastrophe factors (see p. 1003). Kinesin-14 is a C-terminal kinesin that includes the *Drosophila* protein Ncd and the yeast protein Kar3. These kinesins generally travel in the opposite direction from the majority of kinesins, toward the minus end instead of the plus end of a microtubule. (B) Freeze-etch electron micrograph of a kinesin molecule with the head domains on the left. (B, courtesy of John Heuser.)

There Are Two Types of Microtubule Motor Proteins: Kinesins and Dyneins

Kinesin is a motor protein that moves along microtubules. It was first identified in the giant axon of the squid, where it carries membrane-enclosed organelles away from the neuronal cell body toward the axon terminal by walking toward the plus end of microtubules. Kinesin is similar structurally to myosin II in having two heavy chains and two light chains per active motor; these form two globular head motor domains and an elongated coiled-coil tail responsible for heavy chain dimerization. Like myosin, kinesin is a member of a large protein superfamily, for which the motor domain is the only common element (Figure 16–58). The yeast *Saccharomyces cerevisiae* has six distinct kinesins. The nematode *C. elegans* has 16 kinesins, and humans have about 45.

There are at least fourteen distinct families in the kinesin superfamily. Most of them have the motor domain at the N-terminus of the heavy chain and walk toward the plus end of the microtubule. A particularly interesting family has the motor domain at the C-terminus and walks in the opposite direction, toward the minus end of the microtubule. Some kinesin heavy chains lack a coiled-coil sequence and seem to function as monomers, analogous to myosin I. Some others are homodimers, and yet others are heterodimers. Members of the kinesin-5 family can self-associate through the tail domain, forming a bipolar motor that slides oppositely oriented microtubules past one another, much as a myosin II thick filament does for actin filaments. Most kinesins carry a binding site in the tail for either a membrane-enclosed organelle or another microtubule. Many of the kinesin superfamily members have specific roles in mitotic and meiotic spindle formation, and in chromosome separation during cell division.

The **dyneins** are a family of minus-end-directed microtubule motors unrelated to the kinesin superfamily. They are composed of two or three heavy chains (that include the motor domain) and a large and variable number of associated intermediate chains and light chains. The dynein family has two major branches (Figure 16–59). The most ancient branch contains the *cytoplasmic dyneins*, which are typically heavy-chain homodimers, with two large motor domains as heads. Cytoplasmic dyneins are probably found in all eucaryotic cells, and they

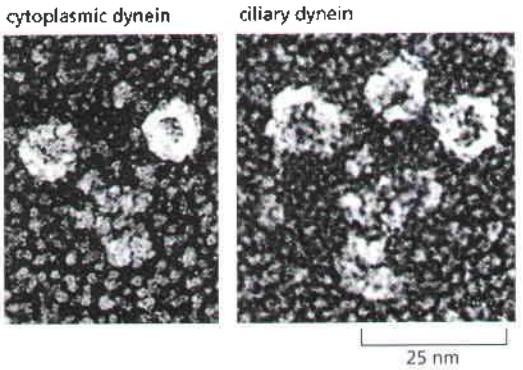


Figure 16–59 Dyneins. Freeze-etch electron micrographs of a molecule of cytoplasmic dynein and a molecule of ciliary (axonemal) dynein. Like myosin II and kinesin, cytoplasmic dynein is a two-headed molecule. The ciliary dynein shown is from a protozoan and has three heads; ciliary dynein from animals has two heads. Note that the dynein head is very large compared with the head of either myosin or kinesin. (Courtesy of John Heuser.)

are important for vesicle trafficking, as well as for localization of the Golgi apparatus near the center of the cell. *Axonemal dyneins*, the other large branch, include heterodimers and heterotrimers, with two or three motor-domain heads, respectively. They are highly specialized for the rapid and efficient sliding movements of microtubules that drive the beating of cilia and flagella (discussed later). A third, minor, branch shares greater sequence similarity with cytoplasmic than with axonemal dyneins but seems to be involved in the beating of cilia.

Dyneins are the largest of the known molecular motors, and they are also among the fastest: axonemal dyneins can move microtubules in a test tube at the remarkable rate of 14 μm/sec. In comparison, the fastest kinesins can move their microtubules at about 2–3 μm/sec. We shall discuss how they work below.

The Structural Similarity of Myosin and Kinesin Indicates a Common Evolutionary Origin

The motor domain of myosins is substantially larger than that of kinesins, about 850 amino acids compared with about 350. The two classes of motor proteins track along different filaments and have different kinetic properties, and they have no identifiable amino acid sequence similarities. However, determination of the three-dimensional structure of the motor domains of myosin and kinesin has revealed that these two motor domains are built around nearly identical cores (**Figure 16–60**). The central force-generating element that the two types of motor proteins have in common includes the site of ATP binding and the machinery necessary to translate ATP hydrolysis into an allosteric conformational change. Large loops extending outward from the central core cause the

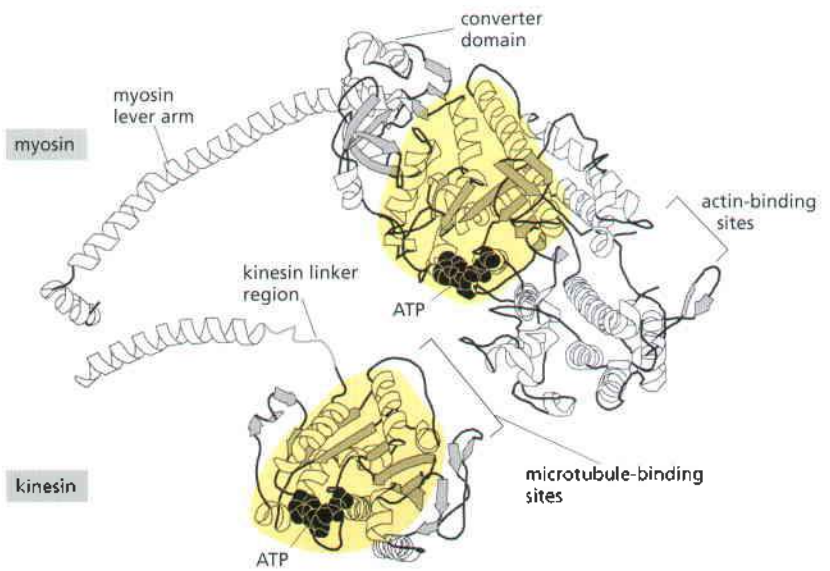


Figure 16–60 X-ray crystal structures of myosin and kinesin heads. The central nucleotide-binding domains of myosin and kinesin (shaded in *yellow*) are structurally very similar. The very different sizes and functions of the two motors are due to major differences in the polymer-binding and force-transduction portions of the motor domain. (Adapted from L.A. Amos and R.A. Cross, *Curr. Opin. Struct. Biol.* 7:239–246, 1997. With permission from Elsevier.)

difference in domain size and are responsible for the choice of track. These loops include the actin-binding and microtubule-binding sites in the myosin and kinesin, respectively. It is thought that both the kinesins and the myosins are descended from a common ancestral motor protein precursor, and that their various specialized functions arose via gene duplication and modification through evolution of the loops coming out from the central core.

An important clue to how the central core is involved in force generation has come from the observation that the motor core also bears some structural resemblance to the nucleotide binding site of the small GTPases of the Ras superfamily. As discussed in Chapter 3 (see Figure 3–72), these proteins exhibit distinct conformations in their GTP-bound (active) and GDP-bound (inactive) forms: mobile "switch" loops in the nucleotide-binding site are in close contact with the γ-phosphate in the GTP-bound state, but these loops swing out when the hydrolyzed γ-phosphate (the terminal phosphate) is released. Although the details of the movement differ for the two motor proteins, and ATP rather than GTP is hydrolyzed, a relatively small structural change in the active site—the presence or absence of a terminal phosphate—is similarly amplified to cause a rotation of a different part of the protein.

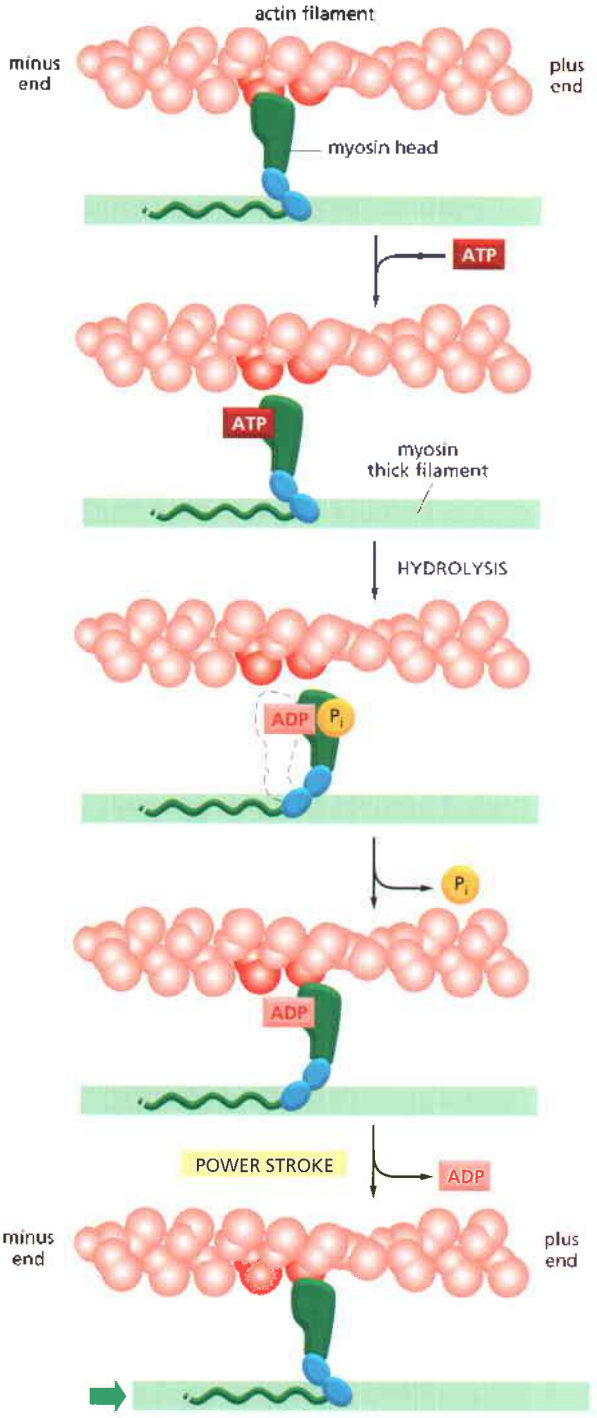
In kinesin and myosin, a switch loop interacts extensively with those regions of the protein involved in microtubule and actin binding, respectively, allowing the structural transitions caused by the ATP hydrolysis cycle to be relayed to the polymer-binding interface. The relay of structural changes between the polymer-binding site and the nucleotide hydrolysis site seems to work in both directions, since the ATPase activity of motor proteins is strongly activated by binding to their filament tracks.

Motor Proteins Generate Force by Coupling ATP Hydrolysis to Conformational Changes

Although the cytoskeletal motor proteins and GTP-binding proteins both use structural changes in their nucleoside-triphosphate-binding sites to produce cyclic interactions with a partner protein, the motor proteins have a further requirement: each cycle of binding and release must propel them forward in a single direction along a filament to a new binding site on the filament. For such unidirectional motion, a motor protein must use the energy derived from ATP binding and hydrolysis to force a large movement in part of the protein molecule. For myosin, each step of the movement along actin is generated by the swinging of an 8.5-nm-long α helix, or *lever arm*, which is structurally stabilized by the binding of light chains. At the base of this lever arm next to the head, there is a piston-like helix that connects movements at the ATP-binding cleft in the head to small rotations of the so-called converter domain (see Figure 16–60). A small change at this point can swing the helix like a long lever, causing the far end of the helix to move by about 5.0 nm.

These changes in the conformation of the myosin are coupled to changes in its binding affinity for actin, allowing the myosin head to release its grip on the actin filament at one point and snatch hold of it again at another. The full mechanochemical cycle of nucleotide binding, nucleotide hydrolysis, and phosphate release (which causes the "power stroke") produces a single step of movement (**Figure 16–61**).

In kinesin, instead of the rocking of a lever arm, the small movements of switch loops at the nucleotide-binding site regulate the docking and undocking of the motor head domain to a long linker region that connects this motor head at one end to the coiled-coil dimerization domain at the other end (see Figure 16–61). When the front (leading) kinesin head is bound to a microtubule before the power stroke, its linker region is relatively unstructured. On the binding of ATP to this bound head, its linker region docks along the side of the head; this throws the second head forward to a position where it will be able to bind a new attachment site on the protofilament, 8 nm closer to the microtubule plus end than the binding site for the first head. The nucleotide hydrolysis cycles in the two heads are closely coordinated, so that this cycle of linker docking and undocking allows



ATTACHED At the start of the cycle shown in this figure, a myosin head lacking a bound nucleotide is locked tightly onto an actin filament in a *rigor* configuration (so named because it is responsible for *rigor mortis*, the rigidity of death). In an actively contracting muscle, this state is very short-lived, being rapidly terminated by the binding of a molecule of ATP.

RELEASED A molecule of ATP binds to the large cleft on the "back" of the head (that is, on the side furthest from the actin filament) and immediately causes a slight change in the conformation of the domains that make up the actin-binding site. This reduces the affinity of the head for actin and allows it to move along the filament. (The space drawn here between the head and actin emphasizes this change, although in reality the head probably remains very close to the actin.)

COCKED The cleft closes like a clam shell around the ATP molecule, triggering a large shape change that causes the head to be displaced along the filament by a distance of about 5 nm. Hydrolysis of ATP occurs, but the ADP and inorganic phosphate (P_i) produced remain tightly bound to the protein.

FORCE-GENERATING A weak binding of the myosin head to a new site on the actin filament causes release of the inorganic phosphate produced by ATP hydrolysis, concomitantly with the tight binding of the head to actin. This release triggers the power stroke—the force-generating change in shape during which the head regains its original conformation. In the course of the power stroke, the head loses its bound ADP, thereby returning to the start of a new cycle.

ATTACHED At the end of the cycle, the myosin head is again locked tightly to the actin filament in a rigor configuration. Note that the head has moved to a new position on the actin filament.

Figure 16–61 The cycle of structural changes used by myosin II to walk along an actin filament. <ATAT> In the myosin II cycle, the head remains bound to the actin filament for only about 5% of the entire cycle time, allowing many myosins to work together to move a single actin filament. (Based on I. Rayment et al., *Science* 261:50–58, 1993. With permission from AAAS.)

the two-headed motor to move in a hand-over-hand (or head-over-head) stepwise manner (**Figure 16–62**), each time taking a discrete 8-nm step.

The coiled-coil domain seems both to coordinate the mechanochemical cycles of the two heads (motor domains) of the kinesin dimer and to determine its directionality of movement. Recall that whereas most members of the kinesin superfamily, with their motor domains at the N-terminus, move toward the plus end of the microtubule, a few superfamily members have their motor domains at the C-terminus and move toward the minus end. Since the motor domains of these two types of kinesins are essentially identical, how can they move in opposite directions? The answer seems to lie in the way in which the heads are connected. In high-resolution images of forward-walking and backward-walking members of the kinesin superfamily bound to microtubules, the heads that are attached to the microtubule are essentially indistinguishable, but the second, unattached heads are oriented differently. This difference in tilt apparently

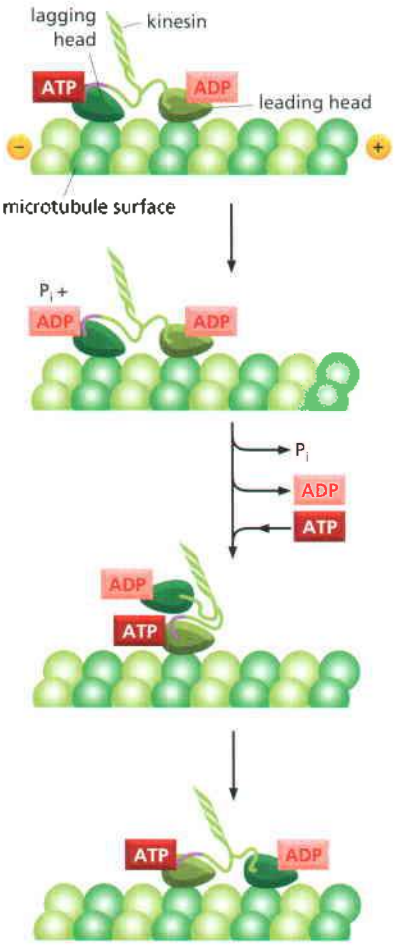


Figure 16–62 The mechanochemical cycle of kinesin. <GAAT> Kinesin-1 is a dimer of two nucleotide-binding motor domains (heads) that are connected through a long coiled-coil tail (see Figure 16–58). The two kinesin motor domains work in a coordinated manner; during a kinesin "step," the rear head detaches from its tubulin binding site, passes the partner motor domain, and then rebinds to the next available tubulin binding site. Using this "hand-over-hand" motion, the kinesin dimer can move for long distances on the microtubule without completely letting go of its track.

At the start of each step, one of the two kinesin heads, the rear or lagging head (*dark green*), is tightly bound to the microtubule and to ATP, while the front or leading head is loosely bound to the microtubule with ADP in its binding site. The forward displacement of the rear motor domain is driven by an exchange of ATP for ADP in the front motor domain (between panels 2 and 3 in this drawing). The binding of ATP to this motor domain causes a small peptide called the "neck linker" to shift from a rearward-pointing to a forward-pointing conformation (the neck linker is drawn here as a connecting line between the motor domain and the intertwined coiled coil). This shift pulls the rear motor domain forward, once it has detached from the microtubule with ADP bound (detachment requires ATP hydrolysis and phosphate (P_i) release). The kinesin molecule is now poised for the next step, which proceeds by an exact repeat of the process shown.

biases the next binding site for the second head, and thereby determines the directionality of motor movement (**Figure 16–63**).

The dynein motor is structurally unrelated to myosins and kinesins, but still follows the general rule of coupling the nucleotide hydrolysis to microtubule binding and unbinding as well as to a force-generating conformational change. A giant heavy chain of more than 500,000 daltons forms the basic structure that creates the movement. Its N-terminal portion forms a tail that binds a set of light chains and connects to the other heavy chains in the dynein molecule, while the major portion of the heavy chain is used to form an elaborate, ring-shaped head. The head consists of a planar ring formed from seven domains: six AAA domains plus the heavy-chain C-terminal domain; it is therefore a more complex relative of the hexameric ATPase discussed in Chapter 6 (see Figure 6–91). A hook-

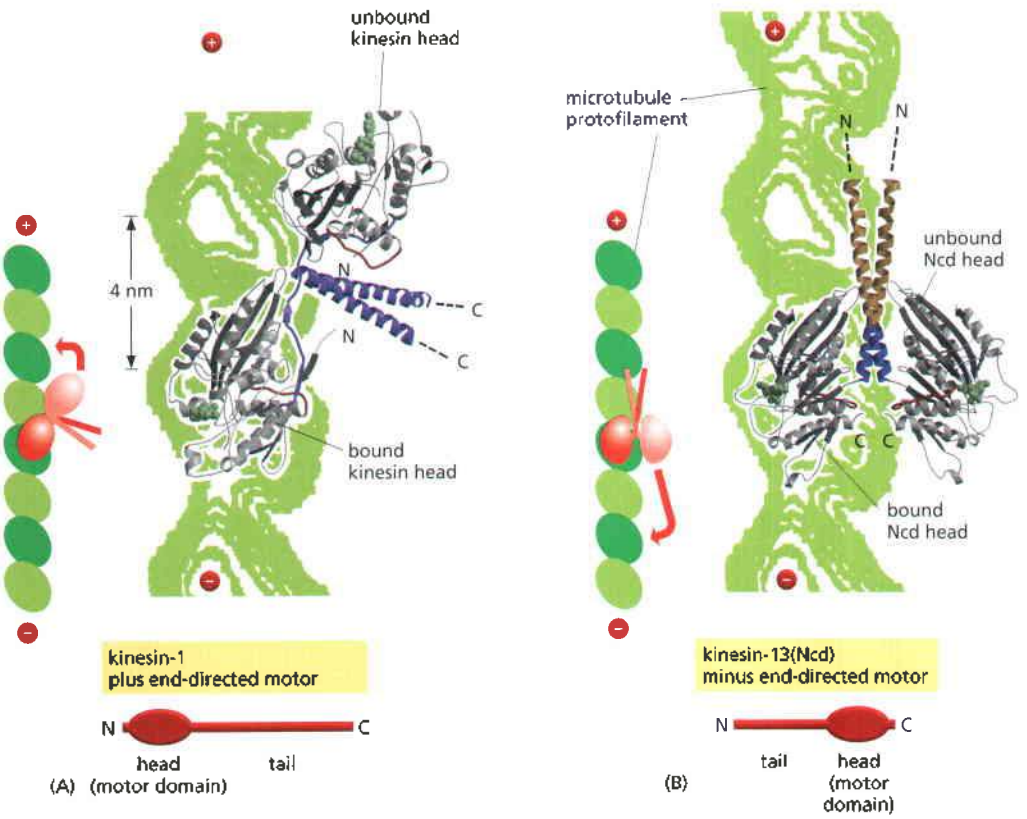


Figure 16–63 Orientation of forward- and backward-walking kinesin superfamily proteins bound to microtubules. These images were generated by fitting the structures of the free motor-protein dimers (determined by x-ray crystallography) onto a lower resolution image of the dimers attached to microtubules (determined by cryoelectron microscopy). (A) Kinesin-1 (conventional kinesin) has its motor domain at the protein's N-terminus and moves toward the plus end of the microtubule. When one head of the dimer is bound to the microtubule in a post-stroke state (with ATP in the nucleotide binding site), the second, unbound head is pointing toward the microtubule plus end, poised to take the next step. (B) Kinesin-13 (called Ncd in *Drosophila*), a minus-end-directed motor with the motor domain at the C-terminus, forms dimers with the opposite orientation. (From E. Sablin et al., *Nature* 395:813–816, 1998. With permission from Macmillan Publishers Ltd.)

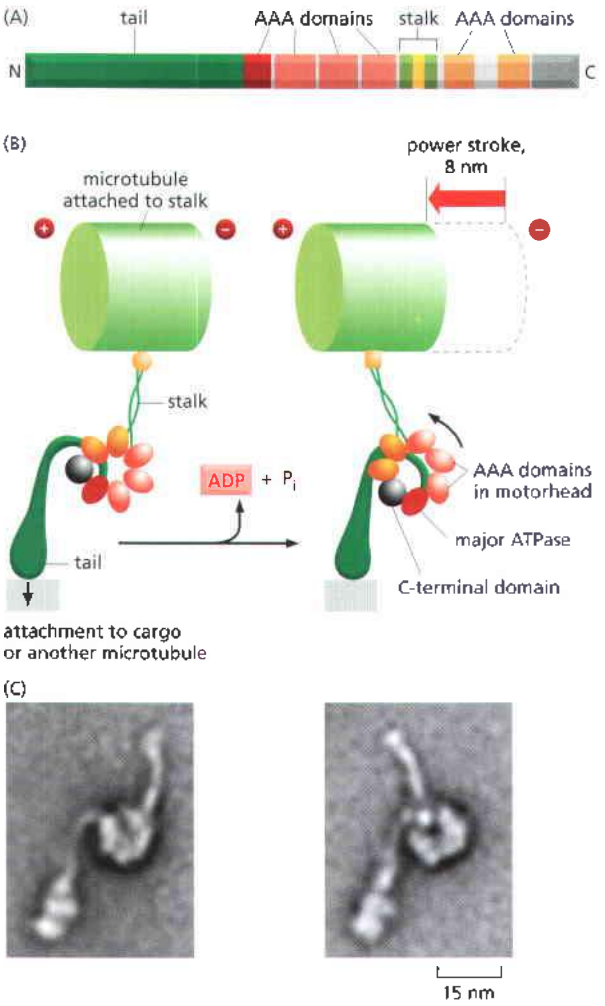


Figure 16–64 The power stroke of dynein. (A) The organization of the domains in each dynein heavy chain. This is a huge molecule, containing nearly 5000 amino acids. The number of heavy chains in a dynein is equal to its number of motor heads. (B) Dynein c is a monomeric flagella dynein found in unicellular green alga *Chlamydomonas reinhardtii*. The large dynein motor head is a planar ring containing a C-terminal domain (*gray*) and six AAA domains, four of which retain ATP-binding sequences, but only one of which (*dark red*) has the major ATPase activity. Extending from the head are a long, coiled-coil stalk with the microtubule binding site at the tip, and a tail with a cargo-attachment site. In the ATP-bound state, the stalk is detached from the microtubule, but ATP hydrolysis causes stalk-microtubule attachment. Subsequent release of ADP and Pi then leads to a large conformational "power stroke" involving rotation of the head and stalk relative to the tail. Each cycle generates a step of about 8 nm along the microtubule towards its minus end. (C) Electron micrographs of purified dyneins in two different conformations representing different steps in the mechanochemical cycle. (B, from S.A. Burgess et al., *Nature* 421:715–718, 2003. With permission from Macmillan Publishers Ltd.)

shaped linker region connects the heavy-chain tail to the AAA domain that is most active as an ATPase. Between the fourth and fifth AAA domains is a heavy chain domain that forms a long anti-parallel coiled-coil stalk. This stalk extends from the top of the ring, with an ATP hydrolysis-regulated microtubule-binding site at its tip. Dynein's "power stroke" is driven by the release of ADP and inorganic phosphate, and it causes the ring to rotate relative to the tail (**Figure 16–64**).

Although kinesin, myosin, and dynein all undergo analogous mechanochemical cycles, the exact nature of the coupling between the mechanical and chemical cycles differs in the three cases. For example, myosin without any nucleotide is tightly bound to its actin track, in a so-called "rigor" state, and it is released from this track by the association of ATP. In contrast, kinesin forms a rigor-like tight association with a microtubule when ATP is bound to the kinesin, and it is hydrolysis of ATP that promotes release of the motor from its track. The mechanochemical cycle of dynein is more similar to myosin than to kinesin, in that nucleotide-free dynein is tightly bound to the microtubule and it is released by binding ATP. However, for dynein the inorganic phosphate and ADP appear to be released at the same time, causing the conformational change driving the power stroke, while for myosin the phosphate is released first and the power stroke does not occur until the ADP subsequently dissociates from the motor head.

Thus, cytoskeletal motor proteins work in a manner highly analogous to GTP-binding proteins, except that in motor proteins the small protein conformational changes (a few tenths of a nanometer) associated with nucleotide hydrolysis are amplified by special protein domains—the lever arm in the case of myosin, the linker in the case of kinesin, and the ring and stalk in the case of dynein—to generate large-scale (several nanometers) conformational changes that move the motor proteins stepwise along their filament tracks. The analogy

between the GTPases and the cytoskeletal motor proteins has recently been extended by the observation that one of the GTP-binding proteins—the bacterial elongation factor G—translates the chemical energy of GTP hydrolysis into directional movement of the mRNA molecule on the ribosome.

Motor Protein Kinetics Are Adapted to Cell Functions

The motor proteins in the myosin and kinesin superfamilies exhibit a remarkable diversity of motile properties, well beyond their choice of different polymer tracks. Most strikingly, a single dimer of kinesin-1 moves in a highly *processive* fashion, traveling for hundreds of ATPase cycles along a microtubule without dissociating. Skeletal muscle myosin II, in contrast, cannot move processively and makes just one or a few steps along an actin filament before letting go. These differences are critical for the motors' various biological roles. A small number of kinesin-1 molecules must be able to transport an organelle all the way down a nerve cell axon, and therefore require a high level of processivity. Skeletal muscle myosin, in contrast, never operates as a single molecule but rather as part of a huge array of myosin II molecules in a thick filament. Here processivity would actually inhibit biological function, since efficient muscle contraction requires that each myosin head perform its power stroke and then quickly get out of the way—in order to avoid interfering with the actions of the other heads attached to the same actin filament.

There are two reasons for the high degree of processivity of kinesin-1 movement. The first is that the mechanochemical cycles of the two motor heads in a kinesin-1 dimer are coordinated with each other, so that one kinesin head does not let go until the other is poised to bind. This coordination allows the motor protein to operate in a hand-over-hand fashion, never allowing the organelle cargo to diffuse away from the microtubule track. In contrast, there is no apparent coordination between the myosin heads in a myosin II dimer. The second reason for the high processivity of kinesin-1 movement is that kinesin-1 spends a relatively large fraction of its ATPase cycle tightly bound to the microtubule. For both kinesin-1 and myosin II, the conformational change that produces the force-generating working stroke must occur while the motor protein is tightly bound to its polymer, and the recovery stroke in preparation for the next step must occur while the motor is unbound. But myosin II spends only about 5% of its ATPase cycle in the tightly bound state, and it is unbound the rest of the time.

What myosin loses in processivity it gains in speed; in an array in which many motor heads are interacting with the same actin filament, a set of linked myosins can move its filament a total distance equivalent to 20 steps during a single cycle time, while kinesins can move only two. Thus, myosin II can typically drive filament sliding much more rapidly than kinesin-1, even though the two different motor proteins hydrolyze ATP at comparable rates and take molecular steps of comparable length. This property is particularly important in the rapid contraction of skeletal muscle, as we will discuss later.

Within each motor protein class, movement speeds vary widely, from about 0.2 to 60 μm/sec for myosins, and from about 0.02 to 2 μm/sec for kinesins. These differences arise from a fine-tuning of the mechanochemical cycle. The number of steps that an individual motor molecule can take in a given time, and thereby the velocity, can be decreased by either decreasing the motor protein's intrinsic ATPase rate or by increasing the proportion of cycle time spent bound to the filament track. For example, myosin V (which acts as a processive vesicle motor) spends up to 90% of its nucleotide cycle tightly bound to the actin filament, in contrast to 5% for myosin II. Moreover, a motor protein can evolve to change the size of each step by either changing the length of the lever arm (for example, the lever arm of myosin V is about three times longer than the lever arm of myosin II) or the angle through which the helix swings (**Figure 16–65**). Each of these parameters varies slightly among different members of the myosin and kinesin families, corresponding to slightly different protein sequences and structures.

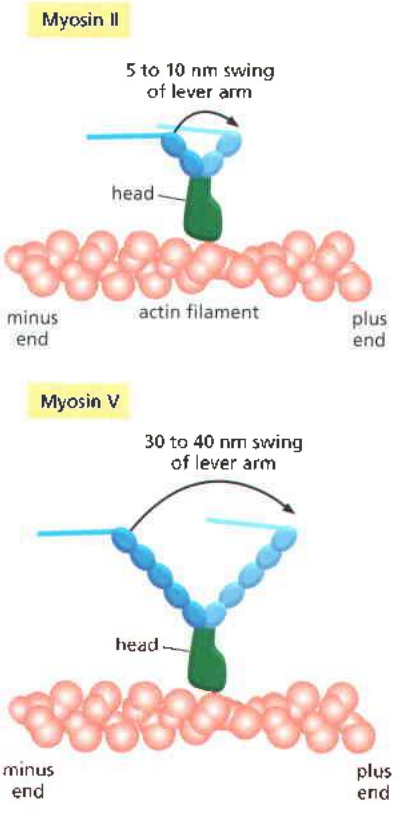


Figure 16–65 **The effect of lever arm length on the step size for a motor protein.** The lever arm of myosin II is much shorter than the lever arm of myosin V. The power stroke in the head swings their lever arms through the same angle, so myosin V is able to take a bigger step than myosin II.

It is assumed that evolution has fine-tuned the behavior of each motor protein, whose function is determined by the identity of the cargo attached through its tail domain, for speed and processivity according to the specific needs of the cell for the function of that particular family member. Whereas there are many different myosin and kinesin family members found in a typical eucaryotic cell, there is usually only one form of cytoplasmic dynein. It is not yet clear how or whether the mechanical properties of cytoplasmic dynein can be modified in response to differing needs of the cell.

Motor Proteins Mediate the Intracellular Transport of Membrane-Enclosed Organelles <CAAT> <AAAT>

A major function of cytoskeletal motors in interphase cells is the transport and positioning of membrane-enclosed organelles. Kinesin was originally identified as the protein responsible for fast axonal transport, the rapid movement of mitochondria, secretory vesicle precursors, and various synapse components down the microtubule highways of the axon to the distant nerve terminals. Although organelles in most cells need not cover such long distances, their polarized transport is equally necessary. A typical microtubule array in an interphase cell is oriented with the minus ends near the center of the cell at the centrosome, and the plus ends extending to the cell periphery. Thus, centripetal movements of organelles or vesicles toward the cell center require the action of minus-end-directed motor proteins such as cytoplasmic dynein, whereas centrifugal movements toward the periphery require plus-end-directed motors such as kinesins.

The clearest example of the effect of microtubules and microtubule motors on the behavior of intracellular membranes is their role in organizing the endoplasmic reticulum (ER) and the Golgi apparatus. The network of ER membrane tubules aligns with microtubules and extends almost to the edge of the cell, whereas the Golgi apparatus is located near the centrosome. When cells are treated with a drug that depolymerizes microtubules, such as colchicine or nocodazole, the ER collapses to the center of the cell, while the Golgi apparatus fragments and disperses throughout the cytoplasm (**Figure 16–66**). *In vitro*, kinesins can tether ER-derived membranes to preformed microtubule tracks, and walk toward the microtubule plus ends, dragging the ER membranes out into tubular protrusions and forming a membranous web very much like the ER in cells. Likewise, in living cells the outward movement of ER tubules toward the cell periphery is associated with microtubule growth. Conversely, dyneins are required for positioning the Golgi apparatus near the cell center, moving Golgi vesicles along microtubule tracks toward minus ends at the centrosome.

The different tails and their associated light chains on specific motor proteins allow the motors to attach to their appropriate organelle cargo. Membrane-associated motor receptors that are sorted to specific membrane-enclosed compartments interact directly or indirectly with the tails of the appropriate kinesin family members. One of these receptors seems to be the amyloid precursor protein, APP, which binds directly to a light chain on the tail of kinesin-1 and is proposed to be a transmembrane motor receptor in nerve-cell axons. The abnormal processing of this protein gives rise to Alzheimer's disease by producing large, stable

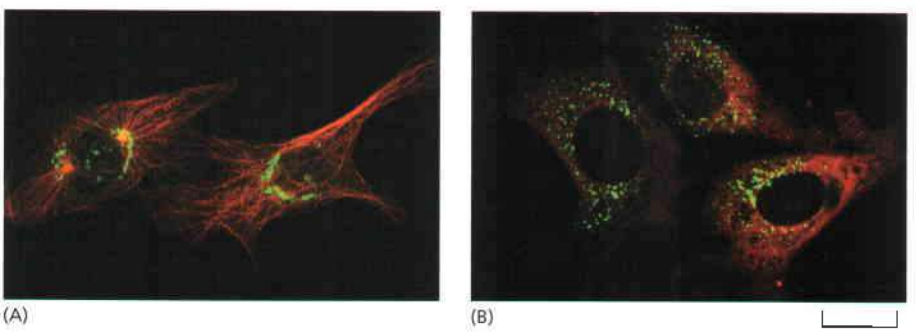


Figure 16–66 **Effect of depolymerizing microtubules on the Golgi apparatus.** (A) In this endothelial cell, the microtubules are labeled in *red*, and the Golgi apparatus is labeled in *green* (using an antibody against a Golgi protein). As long as the system of microtubules remains intact, the Golgi is localized near the centrosome, close to the nucleus at the center of the cell. The cell on the *right* is in interphase, with a single centrosome. The cell on the *left* is in prophase, and the duplicated centrosomes have moved to opposite sides of the nucleus. (B) After exposure to nocodazole, which causes microtubules to depolymerize (see Table 16–2), the Golgi apparatus fragments and is dispersed throughout the cell cytoplasm. (Courtesy of David Shima.)

protein aggregates in nerve cells of the brain (see Figure 6–95). Other receptors for specific kinesins have been identified on the endoplasmic reticulum, as well as on various other membrane-bound organelles that rely on microtubule-based transport for their localization. The JIPs (JNK-interacting proteins), are scaffold proteins associated with cell signaling. These kinesin receptors may provide a link between transport and cell signaling.

For dynein, a large macromolecular assembly often mediates attachment to membranes. Cytoplasmic dynein, itself a huge protein complex, requires association with a second large protein complex called *dynactin* to translocate organelles effectively. The dynactin complex includes a short actin-like filament that forms from the actin-related protein Arp1 (distinct from Arp2 and Arp3, the components of the ARP complex involved in the nucleation of conventional actin filaments). Membranes of the Golgi apparatus are coated with the proteins ankyrin and spectrin, which have been proposed to associate with the Arp1 filament in the dynactin complex to form a planar cytoskeletal array reminiscent of the erythrocyte membrane cytoskeleton (see Figure 10–41). The spectrin array probably gives structural stability to the Golgi membrane, and—via the Arp1 filament—it may mediate the regulated attachment of dynein to the organelle (**Figure 16–67**). In other cases, cytoplasmic dynein motors may interact directly with their cargo. The cytoplasmic tail of rhodopsin, the light-detecting protein found in the rod cells of the eye, binds directly to one of the dynein light chains, and this interaction is required for normal trafficking of rhodopsin in the rod cell.

Motor proteins also have a significant role in organelle transport along actin filaments. The first myosin shown to mediate organelle motility was myosin V, a two-headed myosin with a large step size (see Figure 16–65). In mice and humans, membrane-enclosed pigment granules, called *melanosomes*, are synthesized in cells called *melanocytes* beneath the skin surface. These melanosomes move out to the ends of dendritic processes in the melanocytes, where they are delivered to the overlying keratinocytes that form the skin (and the fur in mice). Myosin V is associated with the surface of melanosomes (**Figure 16–68**) and is able to mediate their actin-based movement in a test tube. In mice, mutations in the myosin V gene result in a "dilute" phenotype, in which fur color looks faded because the melanosomes are not delivered to the keratinocytes efficiently. Other myosins, including myosin I, are associated with endosomes and a variety of other organelles.

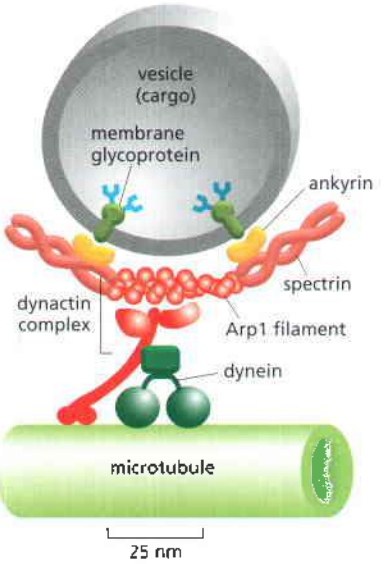


Figure 16–67 A model for the attachment of dynein to a membrane-enclosed organelle. Dynein requires the presence of a large number of accessory proteins to associate with membrane-enclosed organelles. Dynactin is a large complex *(red)* that includes components that bind weakly to microtubules, components that bind to dynein itself, and components that form a small actin-like filament made of the actin-related protein Arp1. It is thought that the Arp1 filament may mediate attachment of this large complex to membrane-enclosed organelles through a network of spectrin and ankyrin, similar to the membrane-associated cytoskeleton of the red blood cell (see Figure 10–41).

The Cytoskeleton Localizes Specific RNA Molecules

In order to concentrate proteins at their site of function, cells often restrict the synthesis of a particular protein by localizing its mRNA molecules, a process that establishes cellular asymmetries. This is particularly important when a parent cell divides to generate two daughters with distinct fates. As another example, several mRNAs encoding proteins involved in synapse function are specifically localized close to synapses in many neurons, and there is evidence that mRNA localization and translation regulation at the synaptic sites play important roles in regulating long-term memory and synaptic plasticity. Not surprisingly, the cytoskeleton and cytoskeletal motor proteins transport and position mRNA molecules in these types of situations.

The giant *Drosophila* oocyte localizes a large number of maternally encoded mRNAs to specific sites within the cell in anticipation of the rapid cell specification events in early embryogenesis (discussed in Chapter 22). A group of mRNAs that encode proteins necessary for proper development of the posterior region of the embryo, including development of the germ cells, is localized posteriorly in the oocyte, and a distinct group of mRNAs encoding proteins necessary for specification of anterior structures in the embryo is localized in the anterior region of the oocyte.

The oocyte takes advantage of its polarized microtubule cytoskeleton, where most microtubule minus ends are clustered in the anterior part of the cell and plus ends near the posterior, to establish these specialized mRNA distributions. For example, the mRNA encoding Bicoid, a transcription factor critical for

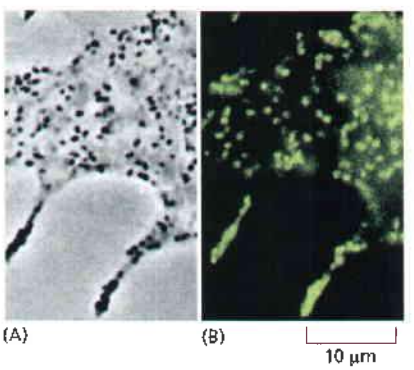
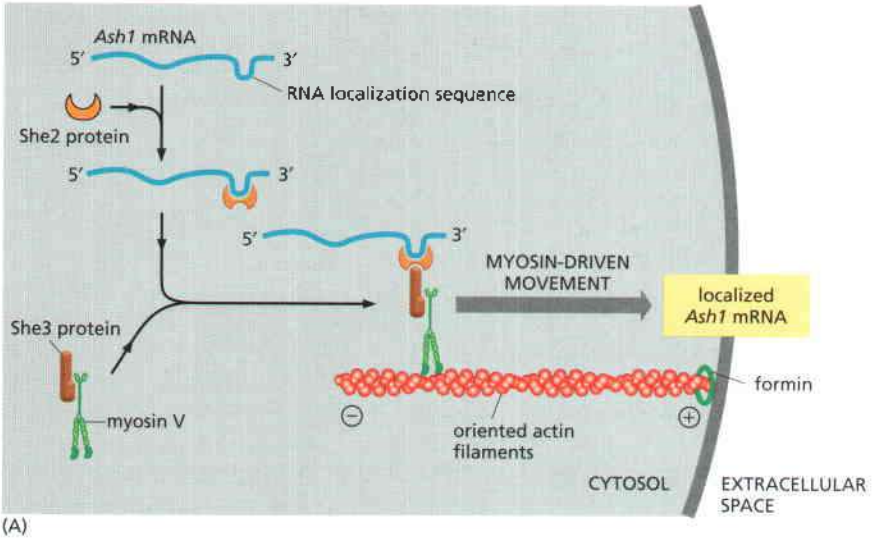


Figure 16–68 Myosin V on melanosomes. (A) Phase-contrast image of a portion of a melanocyte isolated from a mouse. The black spots are melanosomes, which are membrane-enclosed organelles filled with the skin pigment melanin. (B) The same cell labeled with a fluorescent antibody against myosin V. Every melanosome is associated with a large number of copies of this motor protein. (From X. Wu et al., *J. Cell Sci.* 110:847–859, 1997. With permission from The Company of Biologists.)



(A)

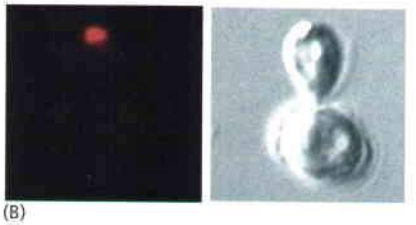

(B)

Figure 16–69 Polarized mRNA localization in the yeast bud tip. (A) The molecular mechanism of *Ash1* mRNA localization, as determined by genetics and biochemistry. (B) Fluorescent *in situ* hybridization (FISH) was used to localize the *Ash1* mRNA (*red*) in this dividing yeast cell. The mRNA is confined to the far tip of the daughter cell (here, still a large bud). Ash1 protein, transcribed from this localized mRNA, is also confined to the daughter cell. (B, courtesy of Peter Takizawa and Ron Vale.)

anterior development, has a structure within the 3′ UTR that binds a protein called Swallow, which in turn binds to a cytoplasmic dynein light chain, presumably enabling its transport to the microtubule minus ends at the cell anterior. Conversely, the transport of mRNA encoding Oskar, a protein necessary for germ cell development in the posterior of the embryo, requires kinesin-1 for its transport to the microtubule plus ends. The anchoring of the mRNAs to their appropriate locations after delivery via microtubules appears to involve the cortical actin cytoskeleton. The mRNA encoding Oskar, for example, binds directly to an actin-binding protein called moesin, a member of the ERM family.

In some cells, mRNA transport as well as anchoring is actin-dependent. The yeast mother and daughter cells retain distinct identities, as revealed by major differences in their subsequent ability to undergo mating-type switching (discussed in Chapter 7) and in the choice of their next bud site. Many of these differences are caused by a gene regulatory protein called *Ash1*. Both *Ash1* mRNA and protein are localized exclusively to the growing bud and therefore end up only in the daughter cell. One of the two type V myosins found in yeast, Myo4p, is required for this asymmetric distribution of *Ash1* mRNA. A genetic screen for other mutations that disrupt the mother/daughter difference has revealed that at least six other gene products that are associated with the cytoskeleton are required for normal polarity; these include one of the formins, tropomyosin, profilin, and actin itself, as well as a complex of two proteins that form a direct link between a specific sequence in the *Ash1* mRNA and the myosin V protein (**Figure 16–69**).

Cells Regulate Motor Protein Function

The cell can regulate the activity of motor proteins and thereby cause either a change in the positioning of its membrane-enclosed organelles or whole-cell movements. Fish melanocytes provide one of the most dramatic examples. These giant cells, which are responsible for rapid changes in skin coloration in several species of fish, contain large pigment granules that can alter their location in response to neuronal or hormonal stimulation (**Figure 16–70**). The pigment granules aggregate or disperse by moving along an extensive network of microtubules. The centrosome nucleates these microtubules, localizing their minus ends in the center of the cell, while the plus ends are distributed around the cell periphery.

The tracking of individual pigment granules (**Figure 16–71**) reveals that the inward movement is rapid and smooth, while the outward movement is jerky, with frequent backward steps. Both the microtubule motors dynein and kinesin are associated with the pigment granules, as well as the actin motor myosin V. The jerky outward movements apparently result from a tug-of-war between the two microtubule motor proteins, with the stronger kinesin winning out overall.

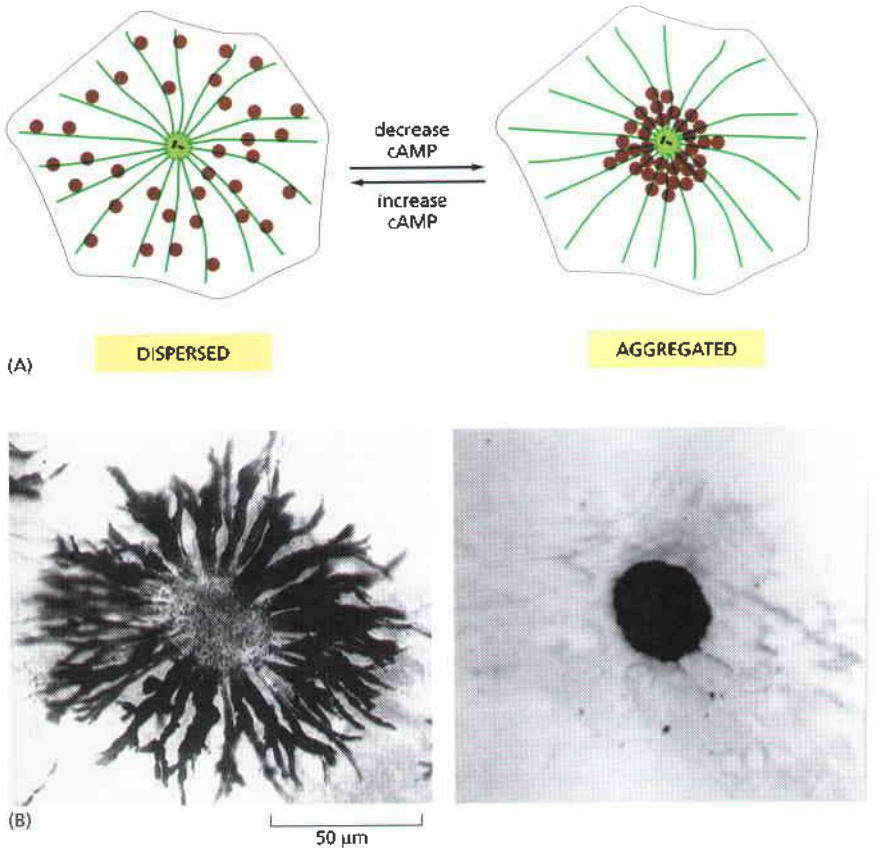


Figure 16–70 Regulated melanosome movements in fish pigment cells. These giant cells, which are responsible for changes in skin coloration in several species of fish, contain large pigment granules, or melanosomes *(brown)*. The melanosomes can change their location in the cell in response to a hormonal or neuronal stimulus. (A) Schematic view of a pigment cell, showing the dispersal and aggregation of melanosomes in response to an increase or decrease in intracellular cyclic AMP (cAMP), respectively. Both redistributions of melanosomes occur along microtubules. (B) Bright-field images of a single cell in a scale of an African cichlid fish, showing its melanosomes either dispersed throughout the cytoplasm *(left)* or aggregated in the center of the cell *(right)*. (B, courtesy of Leah Haimo.)

When the kinesin light chains become phosphorylated after a hormonal stimulation that signals skin color change, kinesin is inactivated, leaving dynein free to drag the pigment granules rapidly toward the cell center, changing the fish's color. In a similar way, the movement of other membrane organelles coated with particular motor proteins is controlled by a complex balance of competing signals that regulate both motor protein attachment and activity.

The cell can also use phosphorylation to regulate myosin activity. In nonmuscle cells, myosin II can be phosphorylated on a variety of sites on both heavy and light chains, affecting both motor activity and thick filament assembly. The myosin II can exist in two different conformational states in such cells, an extended state that can form bipolar filaments, and a bent state in which the tail domain apparently interacts with the motor head. Phosphorylation of the regulatory light chain by the calcium-dependent *myosin light-chain kinase (MLCK)* causes the myosin II to preferentially assume the extended state, which promotes its assembly into a bipolar filament and leads to cell contraction (**Figure 16–72**). MLCK is also activated during mitosis, causing myosin II to assemble into the actin-based contractile ring that pinches the mitotic cell into two. As we will discuss below, myosin phosphorylation is also an important component of the control of contraction in smooth muscle cells. Regulation of other members

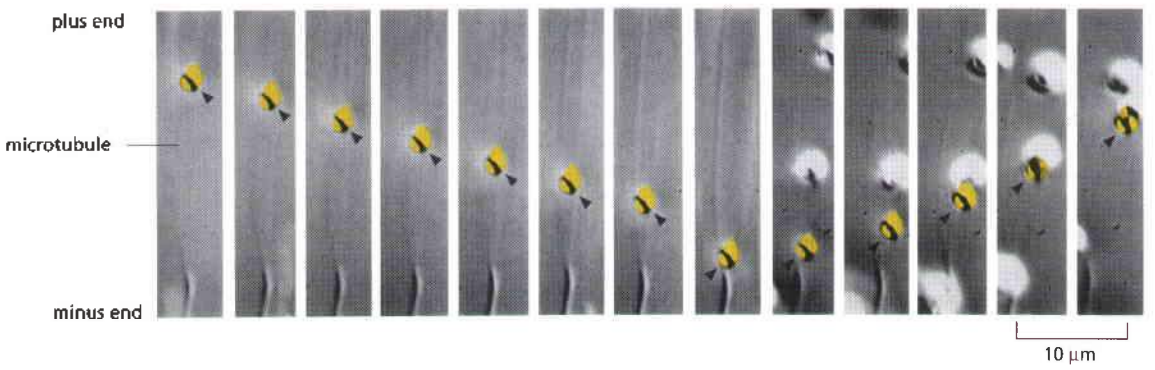


Figure 16–71 Bidirectional movement of a melanosome on a microtubule. An isolated melanosome *(yellow)* moves along a microtubule on a glass slide, from the plus end toward the minus end. Halfway through the video sequence, it abruptly switches direction and moves from the minus end toward the plus end. (From S.L. Rogers et al., *Proc. Natl Acad. Sci. U.S.A.* 94:3720–3725, 1997. With permission from National Academy of Sciences.)

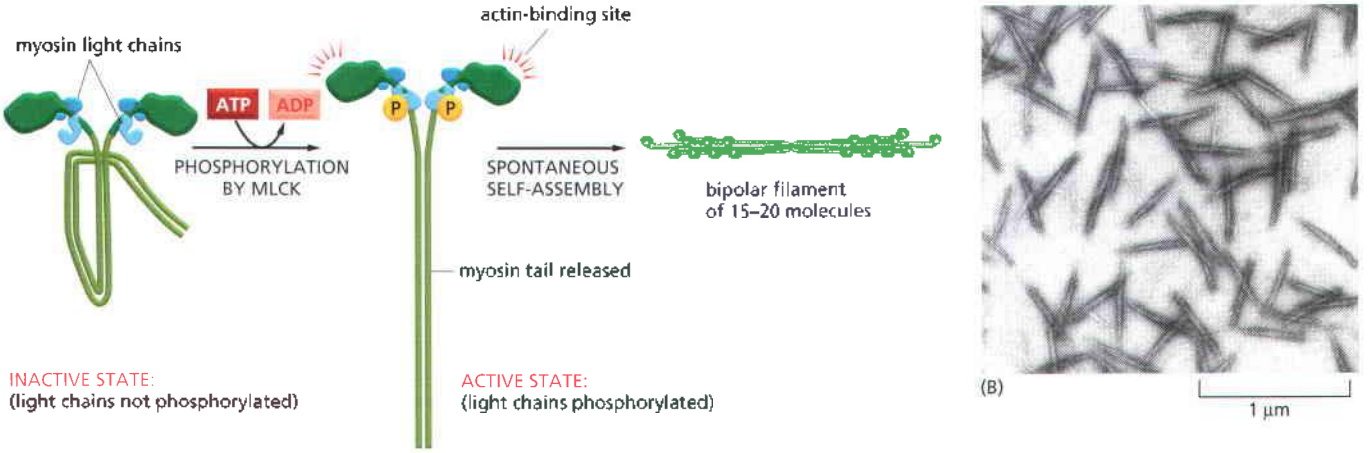


Figure 16–72 **Light-chain phosphorylation and the regulation of the assembly of myosin II into thick filaments.** (A) The controlled phosphorylation by the enzyme myosin light-chain kinase (MLCK) of one of the two light chains (the so-called regulatory light chain, shown in *light blue*) on nonmuscle myosin II in a test tube has at least two effects: it causes a change in the conformation of the myosin head, exposing its actin-binding site, and it releases the myosin tail from a "sticky patch" on the myosin head, thereby allowing the myosin molecules to assemble into short, bipolar, thick filaments. (B) Electron micrograph of negatively stained short filaments of myosin II that have been induced to assemble in a test tube by phosphorylation of their light chains. These myosin II filaments are much smaller than those found in skeletal muscle cells (see Figure 16–55). (B, courtesy of John Kendrick-Jones.)

of the myosin superfamily is not as well understood, but the control of these myosins is likewise thought to involve site-specific phosphorylations.

Summary

Motor proteins use the energy of ATP hydrolysis to move along microtubules or actin filaments. They mediate the sliding of filaments relative to one another and the transport of cargo along filament tracks. All known motor proteins that move on actin filaments are members of the myosin superfamily. The motor proteins that move on microtubules are either members of the kinesin superfamily or the dynein family. The myosin and kinesin superfamilies are diverse, with about 40 genes encoding each type of protein in humans. The only structural element shared among all members of each superfamily is the motor "head" domain. These heads are fused to a wide variety of different "tails," which attach to different types of cargo and enable the various family members to perform different functions in the cell. These functions include the transportation and localization of specific proteins, membrane-enclosed organelles, and mRNAs.

Although myosin and kinesin walk along different tracks and use different mechanisms to produce force and movement by ATP hydrolysis, they share a common structural core, suggesting that they are derived from a common ancestor. The dynein motor protein has independently evolved, and it has a distinct structure and mechanism of action.

THE CYTOSKELETON AND CELL BEHAVIOR

A central challenge in all areas of cell biology is to understand how the functions of many individual molecular components combine to produce complex cell behaviors. The cell behaviors that we describe in this final section all rely on a coordinated deployment of the components and processes that we have explored in the first three sections of the chapter: the dynamic assembly and disassembly of cytoskeletal polymers, the regulation and modification of their structure by polymer-associated proteins, and the actions of motor proteins moving along the polymers. How does the cell coordinate all these activities to define its shape, to enable it to crawl, or to divide it neatly into two at mitosis? These problems of cytoskeletal coordination will challenge scientists for many years to come.

To provide a sense of our present understanding, we first discuss examples where specialized cells build stable arrays of filaments and use highly ordered arrays of motor proteins sliding them relative to each other to generate the large-scale movements of muscle, cilia, and eucaryotic flagella. Next, we consider two important instances where filament dynamics collude with motor protein activity to generate complex, self-organized dynamic structures: the microtubule-based mitotic spindle and the actin arrays involved in cell crawling. Finally, we consider the extraordinary organization and behavior of the neuronal cytoskeleton.

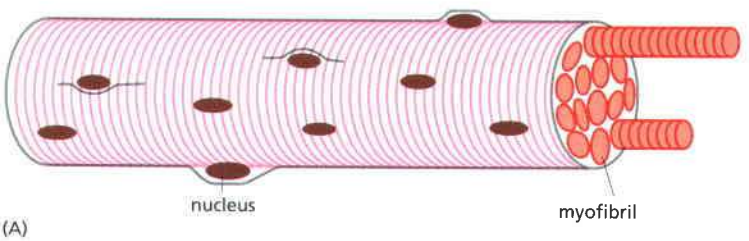
Sliding of Myosin II and Actin Filaments Causes Muscles to Contract

Muscle contraction is the most familiar and the best understood form of movement in animals. In vertebrates, running, walking, swimming, and flying all depend on the rapid contraction of skeletal muscle on its scaffolding of bone, while involuntary movements such as heart pumping and gut peristalsis depend on the contraction of cardiac muscle and smooth muscle, respectively. All these forms of muscle contraction depend on the ATP-driven sliding of highly organized arrays of actin filaments against arrays of myosin II filaments.

Skeletal muscle was a relatively late evolutionary development, and muscle cells are highly specialized for rapid and efficient contraction. The long thin muscle fibers of skeletal muscle are actually huge single cells that form during development by the fusion of many separate cells, as discussed in Chapter 22. The large muscle cell retains the many nuclei of the contributing cells. These nuclei lie just beneath the plasma membrane (**Figure 16–73**). The bulk of the cytoplasm inside is made up of myofibrils, which is the name given to the basic contractile elements of the muscle cell. A **myofibril** is a cylindrical structure 1–2 μm in diameter that is often as long as the giant muscle cell itself. It consists of a long repeated chain of tiny contractile units—called *sarcomeres*, each about 2.2 μm long, which give the vertebrate myofibril its striated appearance (**Figure 16–74**).

Each sarcomere is formed from a miniature, precisely ordered array of parallel and partly overlapping thin and thick filaments. The *thin filaments* are composed of actin and associated proteins, and they are attached at their plus ends to a *Z disc* at each end of the sarcomere. The capped minus ends of the actin filaments extend in toward the middle of the sarcomere, where they overlap with *thick filaments*, the bipolar assemblies formed from specific muscle isoforms of myosin II (see Figure 16–55). When this region of overlap is examined in cross section by electron microscopy, the myosin filaments are seen to be arranged in a regular hexagonal lattice, with the actin filaments evenly spaced between them (**Figure 16–75**). Cardiac muscle and smooth muscle also contain sarcomeres, although the organization is not as regular as that in skeletal muscle.

Sarcomere shortening is caused by the myosin filaments sliding past the actin thin filaments, with no change in the length of either type of filament (Figure 16–74 C and D). Bipolar thick filaments walk toward the plus ends of two sets of thin filaments of opposite orientations, driven by dozens of independent myosin heads that are positioned to interact with each thin filament. Because there is no coordination among the movements of the myosin heads, it is critical



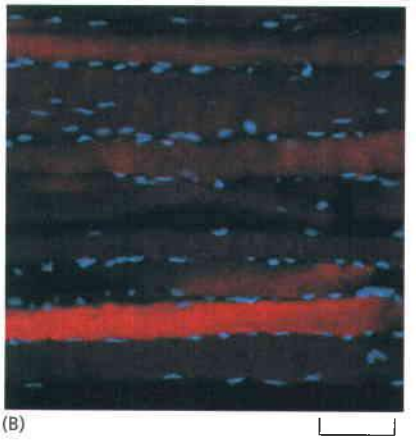


Figure 16–73 Skeletal muscle cells (also called muscle fibers). (A) These huge multinucleated cells form by the fusion of many muscle cell precursors, called myoblasts. In an adult human, a muscle cell is typically 50 μm in diameter and can be up to several centimeters long. (B) Fluorescence micrograph of rat muscle, showing the peripherally located nuclei *(blue)* in these giant cells. Myofibrils are stained *red*; see also Figure 23–46B. (B, courtesy of Nancy L. Kedersha.)

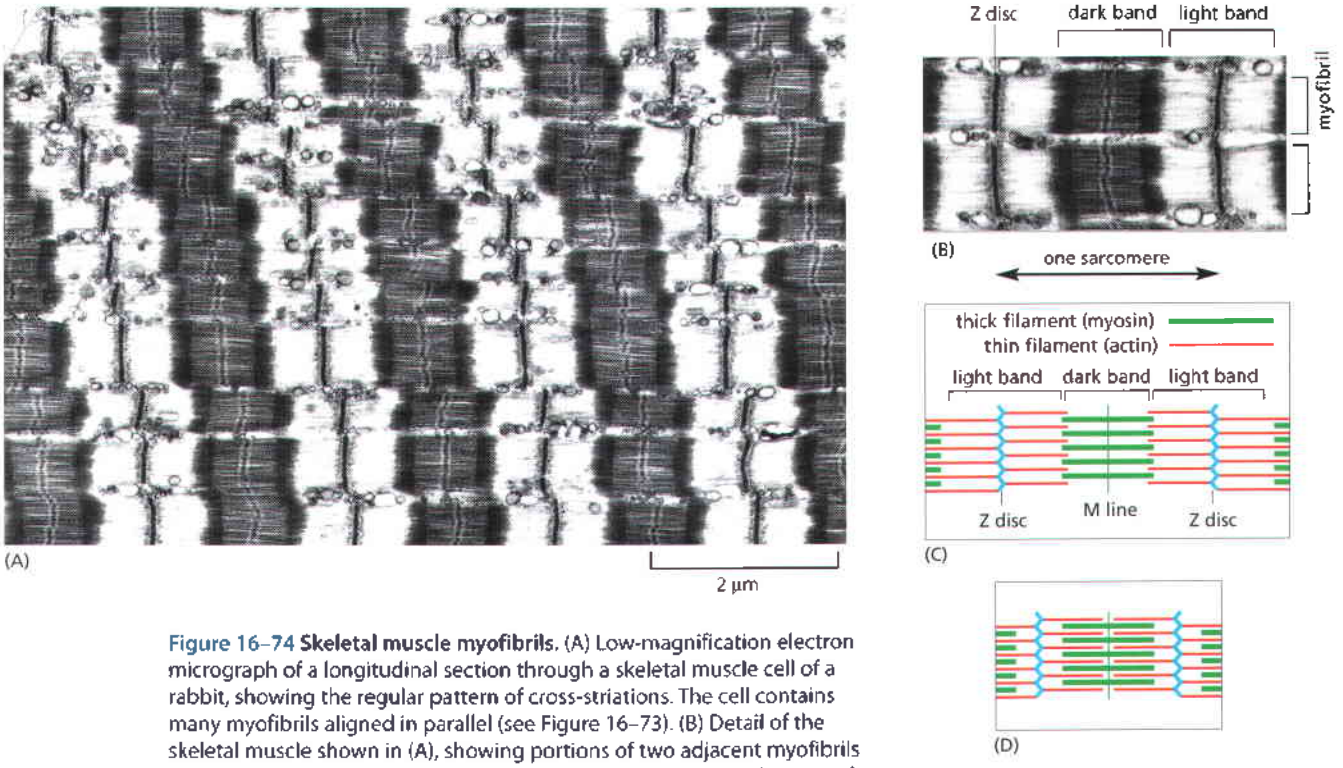


Figure 16–74 Skeletal muscle myofibrils. (A) Low-magnification electron micrograph of a longitudinal section through a skeletal muscle cell of a rabbit, showing the regular pattern of cross-striations. The cell contains many myofibrils aligned in parallel (see Figure 16–73). (B) Detail of the skeletal muscle shown in (A), showing portions of two adjacent myofibrils and the definition of a sarcomere (*black* arrow). (C) Schematic diagram of a single sarcomere, showing the origin of the dark and light bands seen in the electron micrographs. The Z discs, at each end of the sarcomere, are attachment sites for the plus ends of actin filaments (thin filaments); the M line, or midline, is the location of proteins that link adjacent myosin II filaments (thick filaments) to one another. The dark bands, which mark the location of the thick filaments, are sometimes called A bands because they appear anisotropic in polarized light (that is, their refractive index changes with the plane of polarization). The light bands, which contain only thin filaments and therefore have a lower density of protein, are relatively isotropic in polarized light and are sometimes called I bands. (D) When the sarcomere contracts, the actin and myosin filaments slide past one another without shortening. (A and B, courtesy of Roger Craig.)

that they operate with a low processivity, remaining tightly bound to the actin filament for only a small fraction of each ATPase cycle so that they do not hold one another back. Each myosin thick filament has about 300 heads (294 in frog muscle), and each head cycles about five times per second in the course of a

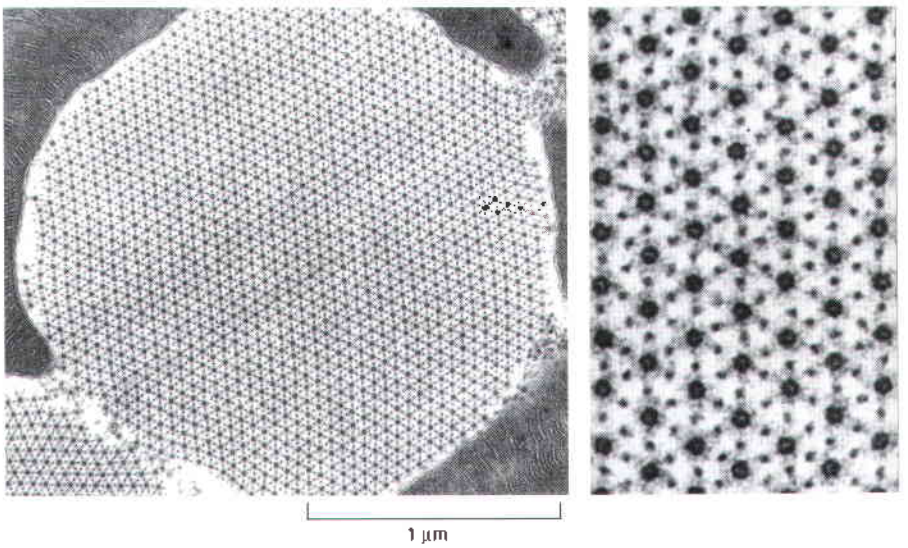
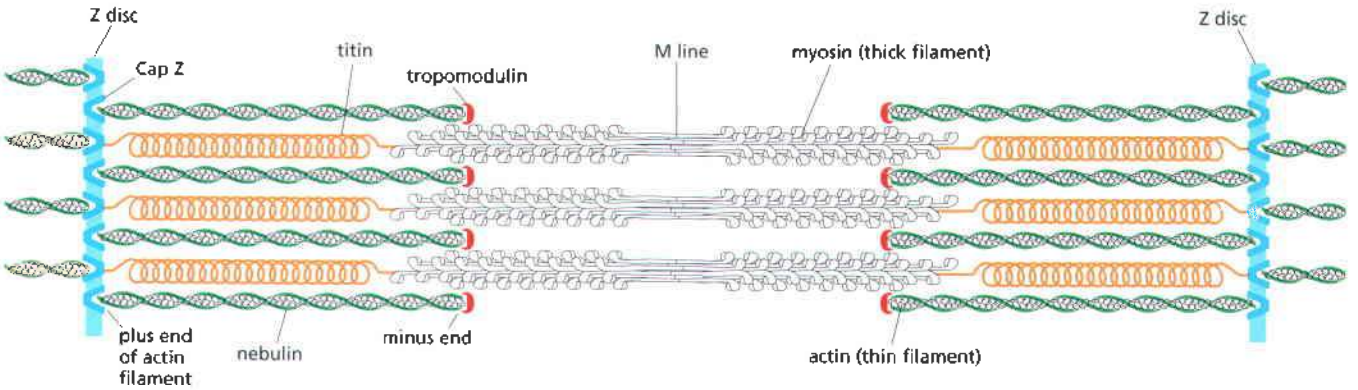


Figure 16–75 Electron micrographs of an insect flight muscle viewed in cross section. The myosin and actin filaments are packed together with almost crystalline regularity. Unlike their vertebrate counterparts, these myosin filaments have a hollow center, as seen in the enlargement on the right. The geometry of the hexagonal lattice is slightly different in vertebrate muscle. (From J. Auber, *J. de Microsc.* 8:197–232, 1969. With permission from Societé française de microscopie électronique.)



rapid contraction—sliding the myosin and actin filaments past one another at rates of up to 15 μm/sec and enabling the sarcomere to shorten by 10% of its length in less than 1/50th of a second. The rapid synchronized shortening of the thousands of sarcomeres lying end-to-end in each myofibril enables skeletal muscle to contract rapidly enough for running and flying, or for playing the piano.

Accessory proteins produce the remarkable uniformity in filament organization, length, and spacing in the sarcomere (**Figure 16–76**). The actin filament plus ends are anchored in the Z disc, which is built from CapZ and α-actinin; the Z disc caps the filaments (preventing depolymerization), while holding them together in a regularly spaced bundle. The precise length of each thin filament is determined by a template protein of enormous size, called *nebulin*, which consists almost entirely of a repeating 35-amino-acid actin-binding motif. Nebulin stretches from the Z disc to the minus end of each thin filament and acts as a "molecular ruler" to dictate the length of the filament. The minus ends of the thin filaments are capped and stabilized by tropomodulin. Although there is some slow exchange of actin subunits at both ends of the muscle thin filament, such that the components of the thin filament turn over with a half-life of several days, the actin filaments in sarcomeres are remarkably stable compared to the dynamic actin filaments characteristic of most other cell types that turn over with half-lives of a few minutes or less.

Opposing pairs of an even longer template protein, called *titin*, position the thick filaments midway between the Z discs. Titin acts as a molecular spring, with a long series of immunoglobulin-like domains that can unfold one by one as stress is applied to the protein. A springlike unfolding and refolding of these domains keeps the thick filaments poised in the middle of the sarcomere and allows the muscle fiber to recover after being overstretched. In *C. elegans*, whose sarcomeres are longer than those in vertebrates, titin is also longer, suggesting that it too serves as a molecular ruler, determining in this case the overall length of each sarcomere (see Figure 3–33).

Figure 16–76 Organization of accessory proteins in a sarcomere. <CTGC> Each giant titin molecule extends from the Z disc to the M line—a distance of over 1 μm. Part of each titin molecule is closely associated with a myosin thick filament (which switches polarity at the M line); the rest of the titin molecule is elastic and changes length as the sarcomere contracts and relaxes. Each nebulin molecule is exactly the length of a thin filament. The actin filaments are also coated with tropomyosin and troponin (not shown; see Figure 16–78) and are capped at both ends. Tropomodulin caps the minus end of the actin filaments, and CapZ anchors the plus end at the Z disc, which also contains α-actinin.

A Sudden Rise in Cytosolic Ca^{2+} Concentration Initiates Muscle Contraction <CTGC>

The force-generating molecular interaction between myosin thick filaments and actin thin filaments takes place only when a signal passes to the skeletal muscle from its motor nerve. Immediately upon arrival of the signal, the muscle cell needs to be able to contract very rapidly, with all the sarcomeres shortening simultaneously. Two major features of the muscle cell are required for extremely rapid contraction. First, as previously discussed, the individual myosin motor heads in each thick filament spend only a small fraction of the ATP cycle time bound to the filament and actively generating force, so many myosin heads can act in rapid succession on the same thin filament without interfering with one another. Second, a specialized membrane system relays the incoming signal rapidly throughout the entire cell. The signal from the nerve triggers an action potential in the muscle cell plasma membrane (discussed in Chapter 11), and

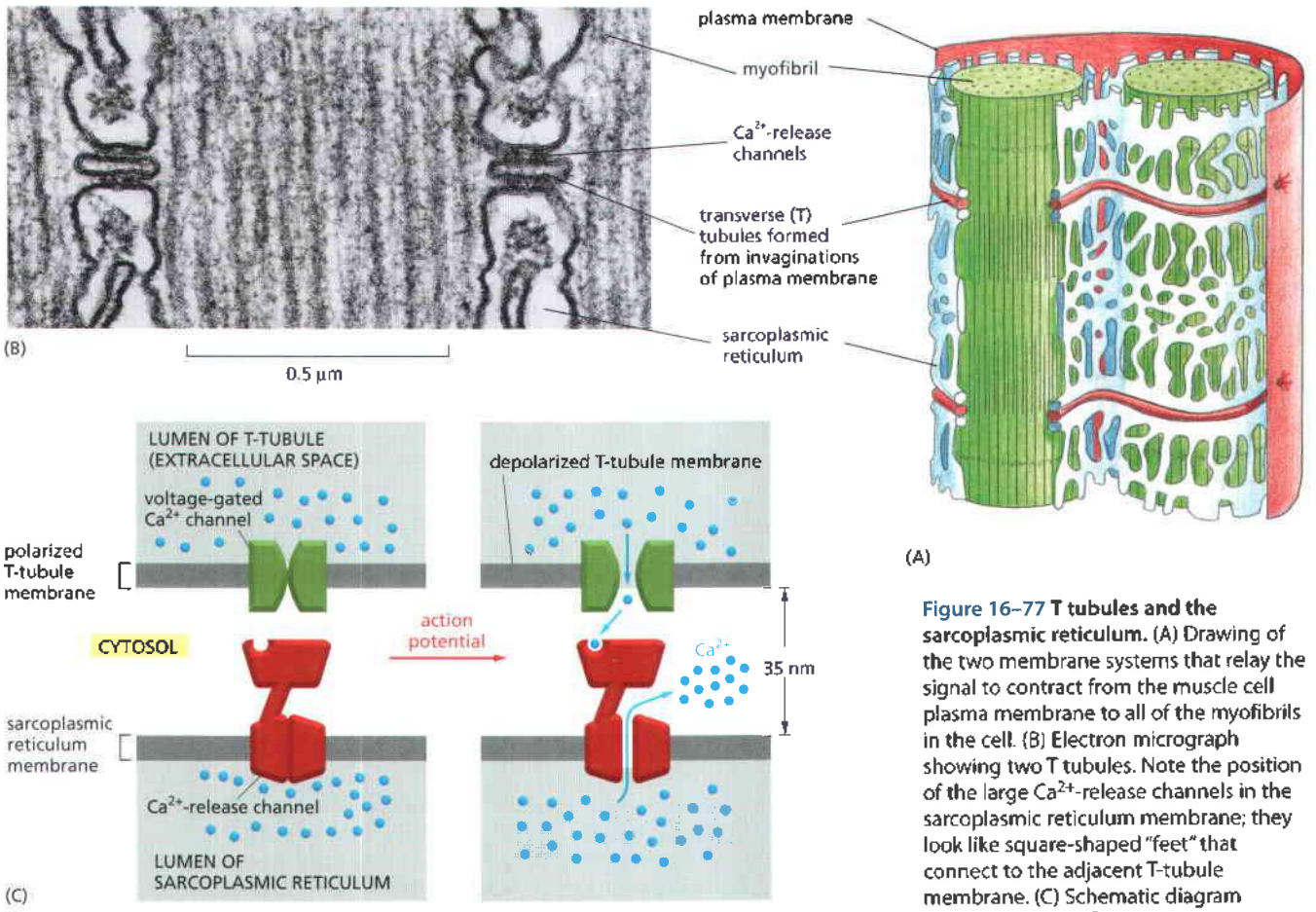


Figure 16–77 T tubules and the sarcoplasmic reticulum. (A) Drawing of the two membrane systems that relay the signal to contract from the muscle cell plasma membrane to all of the myofibrils in the cell. (B) Electron micrograph showing two T tubules. Note the position of the large Ca^{2+}-release channels in the sarcoplasmic reticulum membrane; they look like square-shaped "feet" that connect to the adjacent T-tubule membrane. (C) Schematic diagram showing how a Ca^{2+}-release channel in the sarcoplasmic reticulum membrane is thought to be opened by the activation of a voltage-gated Ca^{2+} channel. (B, courtesy of Clara Franzini-Armstrong.)

this electrical excitation spreads rapidly into a series of membraneous folds, the transverse tubules, or *T tubules*, that extend inward from the plasma membrane around each myofibril. The signal is then relayed across a small gap to the *sarcoplasmic reticulum*, an adjacent web-like sheath of modified endoplasmic reticulum that surrounds each myofibril like a net stocking (Figure 16–77A and B).

When the incoming action potential activates a Ca^{2+} channel in the T-tubule membrane, a Ca^{2+} influx triggers the opening of Ca^{2+}-release channels in the sarcoplasmic reticulum (Figure 16–77C). Ca^{2+} flooding into the cytosol then initiates the contraction of each myofibril. Because the signal from the muscle-cell plasma membrane is passed within milliseconds (via the T tubules and sarcoplasmic reticulum) to every sarcomere in the cell, all of the myofibrils in the cell contract at once. The increase in Ca^{2+} concentration is transient because the Ca^{2+} is rapidly pumped back into the sarcoplasmic reticulum by an abundant, ATP-dependent Ca^{2+}-pump (also called a Ca^{2+}-ATPase) in its membrane (see Figure 11–13). Typically, the cytoplasmic Ca^{2+} concentration is restored to resting levels within 30 msec, allowing the myofibrils to relax. Thus, muscle contraction depends on two processes that consume enormous amounts of ATP: filament sliding, driven by the ATPase of the myosin motor domain, and Ca^{2+} pumping, driven by the Ca^{2+}-pump.

The Ca^{2+} dependence of vertebrate skeletal muscle contraction, and hence its dependence on motor commands transmitted via nerves, is due entirely to a set of specialized accessory proteins that are closely associated with the actin thin filaments. One of these accessory proteins is a muscle form of *tropomyosin*, an elongated molecule that binds along the groove of the actin helix. The other is *troponin*, a complex of three polypeptides, troponins T, I, and C (named for their tropomyosin-binding, inhibitory, and Ca^{2+}-binding activities, respectively). Troponin I binds to actin as well as to troponin T. In a resting muscle, the troponin I–T complex pulls the tropomyosin out of its normal binding groove into a position along the actin filament that interferes with the binding of

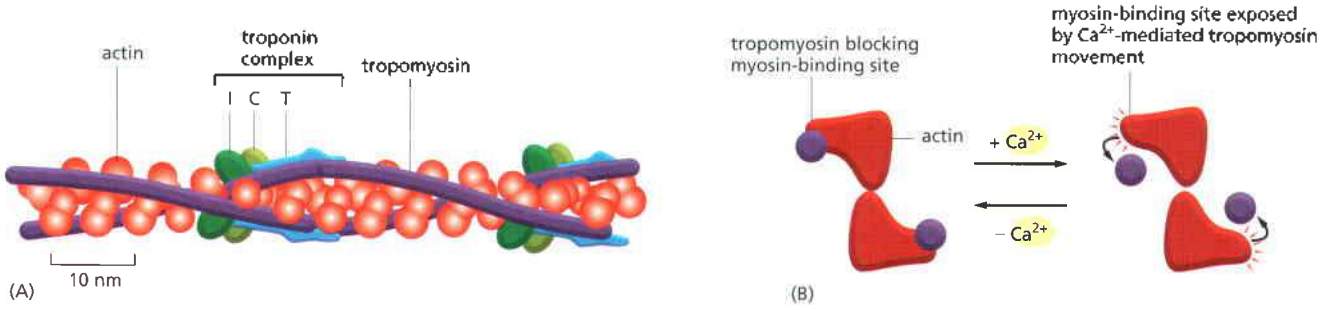


Figure 16–78 The control of skeletal muscle contraction by troponin. (A) A skeletal muscle cell thin filament, showing the positions of tropomyosin and troponin along the actin filament. Each tropomyosin molecule has seven evenly spaced regions with similar amino acid sequences, each of which is thought to bind to an actin subunit in the filament. (B) A thin filament shown end-on, illustrating how Ca^{2+} (binding to troponin) is thought to relieve the tropomyosin blockage of the interaction between actin and the myosin head. (A, adapted from G.N. Phillips, J.P. Fillers and C. Cohen, *J. Mol. Biol.* 192:111–131, 1986. With permission from Academic Press.)

myosin heads, thereby preventing any force-generating interaction. When the level of Ca^{2+} is raised, troponin C—which binds up to four molecules of Ca^{2+}—causes troponin I to release its hold on actin. This allows the tropomyosin molecules to slip back into their normal position so that the myosin heads can walk along the actin filaments (**Figure 16–78**). Troponin C is closely related to the ubiquitous Ca^{2+}-binding protein calmodulin (see Figure 15–44); it can be thought of as a specialized form of calmodulin that has acquired binding sites for troponin I and troponin T, thereby ensuring that the myofibril responds extremely rapidly to an increase in Ca^{2+} concentration.

In smooth muscle cells, so-called because they lack the regular striations of skeletal muscle, contraction is also triggered by an influx of calcium ions, but the regulatory mechanism is different. Smooth muscle forms the contractile portion of the stomach, intestine, and uterus, the walls of arteries, and many other structures requiring slow and sustained contractions. Smooth muscle is composed of sheets of highly elongated spindle-shaped cells, each with a single nucleus. Smooth muscle cells do not express the troponins. Instead, Ca^{2+} influx into the cell regulates contraction by two mechanisms that depend on the ubiquitous calcium binding protein calmodulin.

First, Ca^{2+}-bound calmodulin binds to an actin-binding protein, caldesmon, which blocks the actin sites where the myosin motor heads would normally bind. This causes the caldesmon to fall off of the actin filaments, preparing the filaments for contraction. Second, smooth muscle myosin is phosphorylated on one of its two light chains by myosin light chain kinase (MLCK), as described previously for regulation of nonmuscle myosin II (see Figure 16–72). When the light chain is phosphorylated, the myosin head can interact with actin filaments and cause contraction; when it is dephosphorylated, the myosin head tends to dissociate from actin and becomes inactive (in contrast to nonmuscle myosin II, light chain dephosphorylation does not cause thick filament disassembly in smooth muscle cells). MLCK requires bound Ca^{2+}/calmodulin to be fully active.

External signaling molecules such as adrenaline (epinephrine) can also regulate the contractile activity of smooth muscle. Adrenaline binding to its G-protein-coupled cell surface receptor causes an increase in the intracellular level of cyclic AMP, which in turn activates cyclic-AMP-dependent protein kinase (PKA) (see Figure 15–35). PKA phosphorylates and inactivates MLCK, thereby causing the smooth muscle cell to relax.

The phosphorylation events that regulate contraction in smooth muscle cells occur realtively slowly, so that maximum contraction often requires nearly a second (compared with the few milliseconds required for contraction of a skeletal muscle cell). But rapid activation of contraction is not important in smooth muscle: its myosin II hydrolyzes ATP about 10 times more slowly than skeletal muscle myosin, producing a slow cycle of myosin conformational changes that results in slow contraction.

Heart Muscle Is a Precisely Engineered Machine <AGGT>

The heart is the most heavily worked muscle in the body, contracting about 3 billion (3×10^9) times during the course of a human lifetime. This number is about the same as the average number of revolutions in the lifetime of an automobile's internal combustion engine. Heart cells express several specific isoforms of cardiac muscle myosin and cardiac muscle actin. Even subtle changes in these contractile proteins expressed in the heart—changes that would not cause any noticeable consequences in other tissues—can cause serious heart disease (**Figure 16–79**).

The normal cardiac contractile apparatus is such a highly tuned machine that a tiny abnormality anywhere in the works can be enough to gradually wear it down over years of repetitive motion. *Familial hypertrophic cardiomyopathy* is a frequent cause of sudden death in young athletes. It is a genetically dominant inherited condition that affects about two out of every thousand people, and it is associated with heart enlargement, abnormally small coronary vessels, and disturbances in heart rhythm (cardiac arrhythmias). The cause of this condition is either any one of over 40 subtle point mutations in the genes encoding cardiac β myosin heavy chain (almost all causing changes in or near the motor domain), or one of about a dozen mutations in other genes encoding contractile proteins—including myosin light chains, cardiac troponin, and tropomyosin. Minor missense mutations in the cardiac actin gene cause another type of heart condition, called *dilated cardiomyopathy*, that also frequently results in early heart failure.



Figure 16–79 **Effect on the heart of a subtle mutation in cardiac myosin.** *Left*, normal heart from a 6-day old mouse pup. *Right*, heart from a pup with a point mutation in both copies of its cardiac myosin gene, changing Arg 403 to Gln. The arrows indicate the atria. In the heart from the pup with the cardiac myosin mutation, both atria are greatly enlarged (hypertrophic), and the mice die within a few weeks of birth. (From D. Fatkin et al., *J. Clin. Invest.* 103:147, 1999. With permission from The Rockefeller University Press.)

Cilia and Flagella Are Motile Structures Built from Microtubules and Dyneins

Just as myofibrils are highly specialized and efficient motility machines built from actin and myosin filaments, cilia and flagella are highly specialized and efficient motility structures built from microtubules and dynein. Both cilia and flagella are hair-like cell appendages that have a bundle of microtubules at their core. **Flagella** are found on sperm and many protozoa. By their undulating motion, they enable the cells to which they are attached to swim through liquid media (**Figure 16–80A**). **Cilia** tend to be shorter than flagella and are organized in a similar fashion, but they beat with a whip-like motion that resembles the breast stroke in swimming (Figure 16–80B). The cycles of adjacent cilia are almost but not quite in synchrony, creating the wave-like patterns that can be seen in fields of beating cilia under the microscope. Ciliary beating can either propel single cells through a fluid (as in the swimming of the protozoan *Paramecium*) or can move fluid over the surface of a group of cells in a tissue. In the human body, huge numbers of cilia (10^9/cm^2 or more) line our respiratory tract, sweeping layers of mucus, trapped particles of dust, and bacteria up to the mouth where they are swallowed and ultimately eliminated. Likewise, cilia along the oviduct help to sweep eggs toward the uterus.

The movement of a cilium or a flagellum is produced by the bending of its core, which is called the **axoneme**. The axoneme is composed of microtubules and their associated proteins, arranged in a distinctive and regular pattern. Nine special doublet microtubules (comprising one complete and one partial microtubule fused together so that they share a common tubule wall) are arranged in

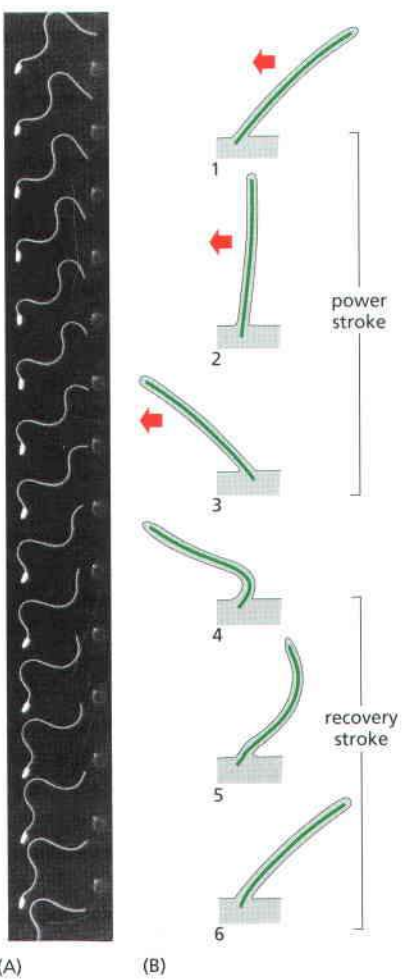


Figure 16–80 **The contrasting motions of flagella and cilia.** (A) The wave-like motion of the flagellum of a sperm cell from a tunicate. The cell was photographed with stroboscopic illumination at 400 flashes per second. Note that waves of constant amplitude move continuously from the base to the tip of the flagellum. (B) The beat of a cilium, which resembles the breast stroke in swimming. A fast power stroke (*red* arrows), in which fluid is driven over the surface of the cell, is followed by a slow recovery stroke. Each cycle typically requires 0.1–0.2 sec and generates a force perpendicular to the axis of the axoneme (the ciliary core). (A, courtesy of C.J. Brokaw.)

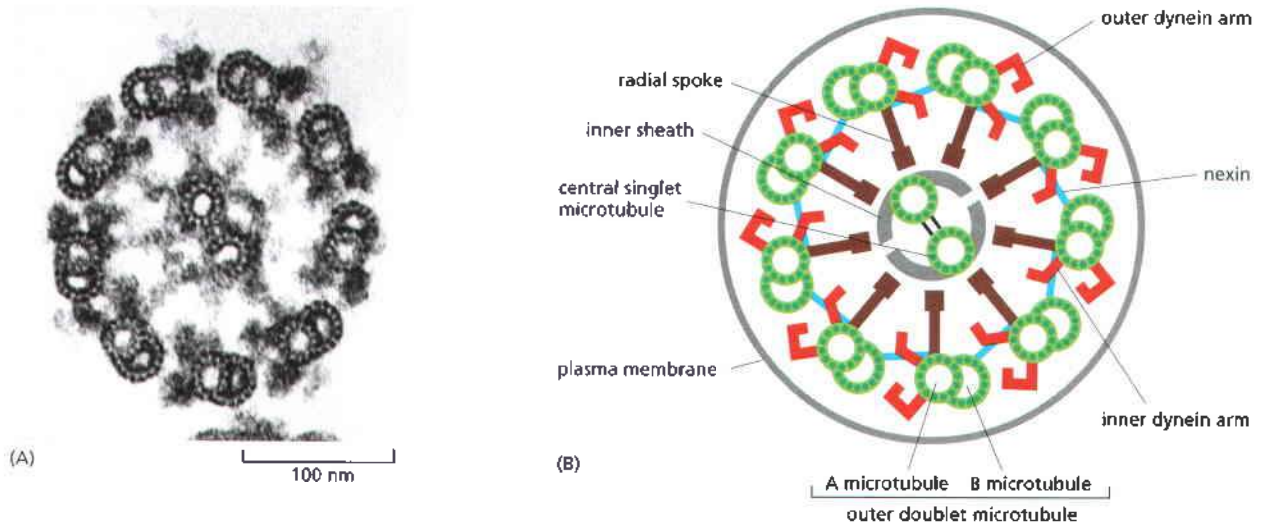


Figure 16–81 **The arrangement of microtubules in a flagellum or cilium.** (A) Electron micrograph of the flagellum of a green-alga cell *(Chlamydomonas)* shown in cross section, illustrating the distinctive "9 + 2" arrangement of microtubules. (B) Diagram of the parts of a flagellum or cilium. The various projections from the microtubules link the microtubules together and occur at regular intervals along the length of the axoneme. (A, courtesy of Lewis Tilney.)

a ring around a pair of single microtubules (**Figure 16–81**). Almost all forms of eucaryotic flagella and cilia (from protozoans to humans) have this characteristic arrangement. The microtubules extend continuously for the length of the axoneme, which can be 10–200 μm. At regular positions along the length of the microtubules, accessory proteins cross-link the microtubules together.

Molecules of *ciliary dynein* form bridges between the neighboring doublet microtubules around the circumference of the axoneme (**Figure 16–82**). When the motor domain of this dynein is activated, the dynein molecules attached to one microtubule doublet (see Figure 16–64) attempt to walk along the adjacent microtubule doublet, tending to force the adjacent doublets to slide relative to one another, much as actin thin filaments slide during muscle contraction. However, the presence of other links between the microtubule doublets prevents this sliding, and the dynein force is instead converted into a bending motion (**Figure 16–83**).

The length of flagella is carefully regulated. If one of the two flagella on a *Chlamydomonas* cell is amputated, the remaining one will transiently shrink as the stump regrows until they reach the same length, and then the two shortened flagella will continue to elongate until both are as long as they were on the unperturbed cell. New flagellar components including tubulin and dynein are incorporated into the growing flagella at the distal tips. Thus, even in these

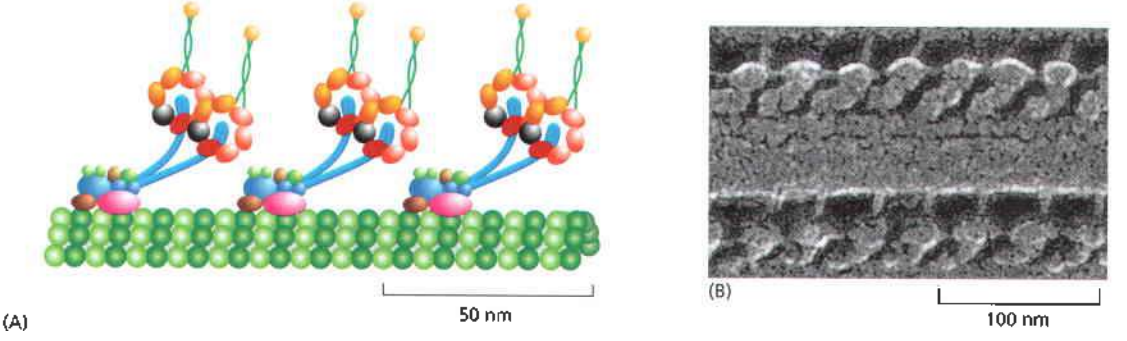


Figure 16–82 **Ciliary dynein.** Ciliary (axonemal) dynein is a large protein assembly (nearly 2 million daltons) composed of 9–12 polypeptide chains, the largest of which is the heavy chain of more than 500,000 daltons. (A) The heavy chains form the major portion of the globular head and stem domains, and many of the smaller chains are clustered around the base of the stem. There are two heads in the outer dynein in metazoans, but three heads in protozoa, each formed from their own heavy chain (see Figure 16–59B for a view of an isolated molecule). The tail of the molecule binds tightly to an A microtubule in an ATP-independent manner, while the large globular heads have an ATP-dependent binding site for a B microtubule (see Figure 16–81). When the heads hydrolyze their bound ATP, they move toward the minus end of the B microtubule, thereby producing a sliding force between the adjacent microtubule doublets in a cilium or flagellum. For details, see Figure 16–64. (B) Freeze-etch electron micrograph of a cilium showing the dynein arms projecting at regular intervals from the doublet microtubules. (B, courtesy of John Heuser.)

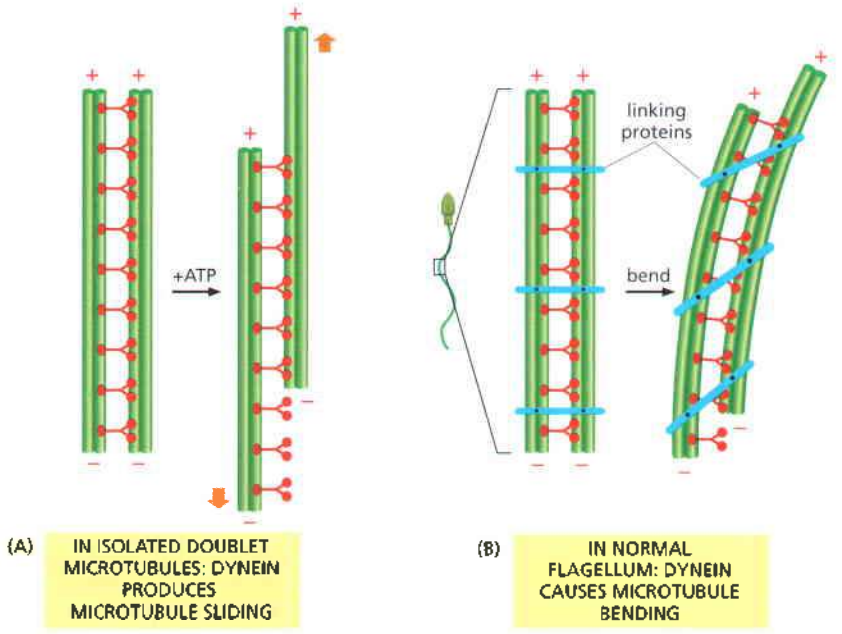


Figure 16–83 The bending of an axoneme. (A) When axonemes are exposed to the proteolytic enzyme trypsin, the linkages holding neighboring doublet microtubules together are broken. In this case, the addition of ATP allows the motor action of the dynein heads to slide one pair of doublet microtubules against the other pair. (B) In an intact axoneme (such as in a sperm), flexible protein links prevent the sliding of the doublet. The motor action therefore causes a bending motion, creating waves or beating motions, as seen in Figure 16–80.

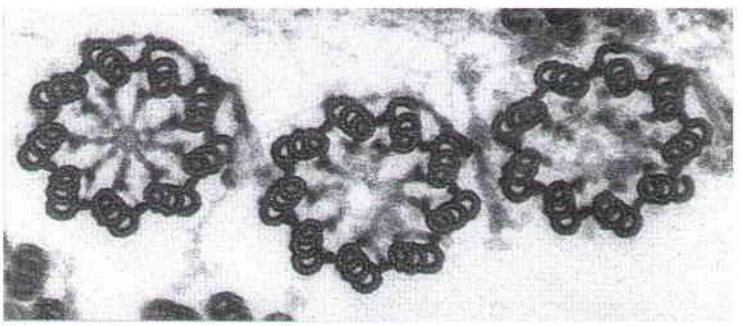
highly ordered and stable filament-motor arrays, cells use the intrinsic flexibility and adaptability of the cytoskeleton to respond rapidly and dynamically to the changes they experience.

In humans, hereditary defects in ciliary dynein cause Kartagener's syndrome. The syndrome is characterized by male sterility due to immotile sperm, a high susceptibility to lung infections owing to the paralyzed cilia in the respiratory tract that fail to clear debris and bacteria, and defects in determination of the left–right axis of the body during early embryonic development (discussed in Chapter 22).

Bacteria also swim using cell surface structures called flagella, but these do not contain microtubules or dynein and do not wave or beat. Instead, *bacterial flagella* are long, rigid helical filaments, made up of repeating subunits of the protein flagellin. The flagella rotate like propellers, driven by a special rotary motor embedded in the bacterial cell wall (see Figure 15–71). The use of the same name to denote these two very different types of swimming apparatus is an unfortunate historical accident.

Structures called *basal bodies* firmly root eucaryotic cilia and flagella at the cell surface. The basal bodies have the same form as the centrioles that are found embedded at the center of animal centrosomes, with nine groups of fused triplet microtubules arranged in a cartwheel (**Figure 16–84**). Indeed, in some organisms, basal bodies and centrioles are functionally interconvertible: during each mitosis in the unicellular alga *Chlamydomonas*, for example, the flagella are resorbed, and the basal bodies move into the cell interior and become part of the spindle poles. New centrioles and basal bodies arise by a curious replication process, in which a smaller daughter is formed perpendicular to the original structure by a still mysterious mechanism (see Figure 17–31).

Figure 16–84 Basal bodies. (A) Electron micrograph of a cross section through three basal bodies in the cortex of a protozoan. (B) Diagram of a basal body viewed from the side. Each basal body forms the lower portion of a ciliary axoneme and is composed of nine sets of triplet microtubules, each triplet containing one complete microtubule (the A microtubule) fused to two incomplete microtubules (the B and C microtubules). Other proteins (shown in *red* in B) form links that hold the cylindrical array of microtubules together. The arrangement of microtubules in a centriole is essentially the same (see Figure 16–31). (A, courtesy of D.T. Woodrow and R.W. Linck.)





Even in animal cells that lack fully developed beating cilia or flagella, centrioles frequently nucleate the growth of a non-motile, microtubule-rich surface projection called a *primary cilium*. Primary cilia are usually only a few micrometers in length and lack dynein. They are found on the surface of many different cell types including fibroblasts, epithelial cells, neurons, bone cells, and chondrocytes (cartilage cells). Many signaling proteins are concentrated in the primary cilium, including proteins involved in the Hedgehog signaling pathway (see p. 950), and receptors for neurotransmitters on neurons in the central nervous system. On kidney epithelial cells, primary cilia act as flow sensors that detect the movement of fluid through the kidney tubules. Mechanosensitive calcium channels are opened when the fluid flow bends the primary cilia, regulating kidney cell growth and proliferation. Loss of the calcium channel or other structural components of the primary cilium in the kidney cells causes polycystic kidney disease, a common genetic disorder that causes overproliferation of the kidney epithelial cells—resulting in the formation of large fluid-filled cysts throughout the organs and eventually in kidney failure. Another specialized kind of primary cilium that is unusual in being able to beat is required for establishing left-right asymmetry in the developing embryo (see Figure 22–87).

Construction of the Mitotic Spindle Requires Microtubule Dynamics and the Interactions of Many Motor Proteins

Myofibrils and cilia are relatively permanent structures specialized to produce repetitive movement. But most cell movements depend on labile structures that appear at specific stages of the cell cycle or in response to external signals and then disappear once they complete their jobs. The most familiar of these are the mitotic spindle and the contractile ring that form during cell division. In Chapter 17, we will describe in detail both the process of mitosis and the cell cycle control system that determines the timing of the events of cell division. Here, we briefly discuss a few of the cytoskeletal mechanisms that contribute to the construction and mechanical function of the mitotic spindle.

The construction of the mitotic spindle is a particularly important and fascinating example of the power of self-organization by teams of motor proteins interacting with dynamic cytoskeletal filaments. It also features the active participation of the chromosomes. In a rapid sequence of events that typically takes less than an hour in animal cells, the interphase array of microtubules is completely disassembled and reorganized to form the bipolar spindle structure that is responsible for segregating the replicated chromosomes with perfect fidelity to the two daughter cells. Because of the central importance of reliability in transmission of the genetic material, the construction and functioning of the mitotic spindle feature a tremendous degree of redundancy, so that if one set of mechanisms fails for any reason, there are backup mechanisms in place to ensure reliable chromosome partitioning.

In early mitosis, there are dramatic changes in the dynamic behavior and average length of the microtubules. In the interphase array the microtubules are typically long and undergo rare catastrophes, but during mitosis the microtubules are shorter and much more dynamic. Microtubule nucleation and assembly are enhanced in the regions around the condensed chromosomes. As microtubules assemble on condensed chromatin pointing in random directions, the coordinated actions of several motor proteins build a coherent bipolar spindle from the disorganized microtubule mass. First, the bipolar kinesin-5 (see Figure 16–58) bundles the microtubules into a parallel array and slides microtubules that are oriented in opposite directions away from each other. Next, another kinesin that is bound to chromosome arms, kinesin-4, walks toward the plus ends of chromosome-associated microtubules and pushes their minus ends away from the chromosome mass. Finally, the minus-end directed motors cytoplasmic dynein and kinesin-14 form oligomeric complexes with scaffold proteins that gather the microtubule minus ends together to form the spindle poles. In most animal cells, these processes are guided by a pair of centrosomes that help to nucleate and organise these microtubule minus ends. The final result is the elegantly balanced bipolar mitotic spindle (**Figure 16–85**).

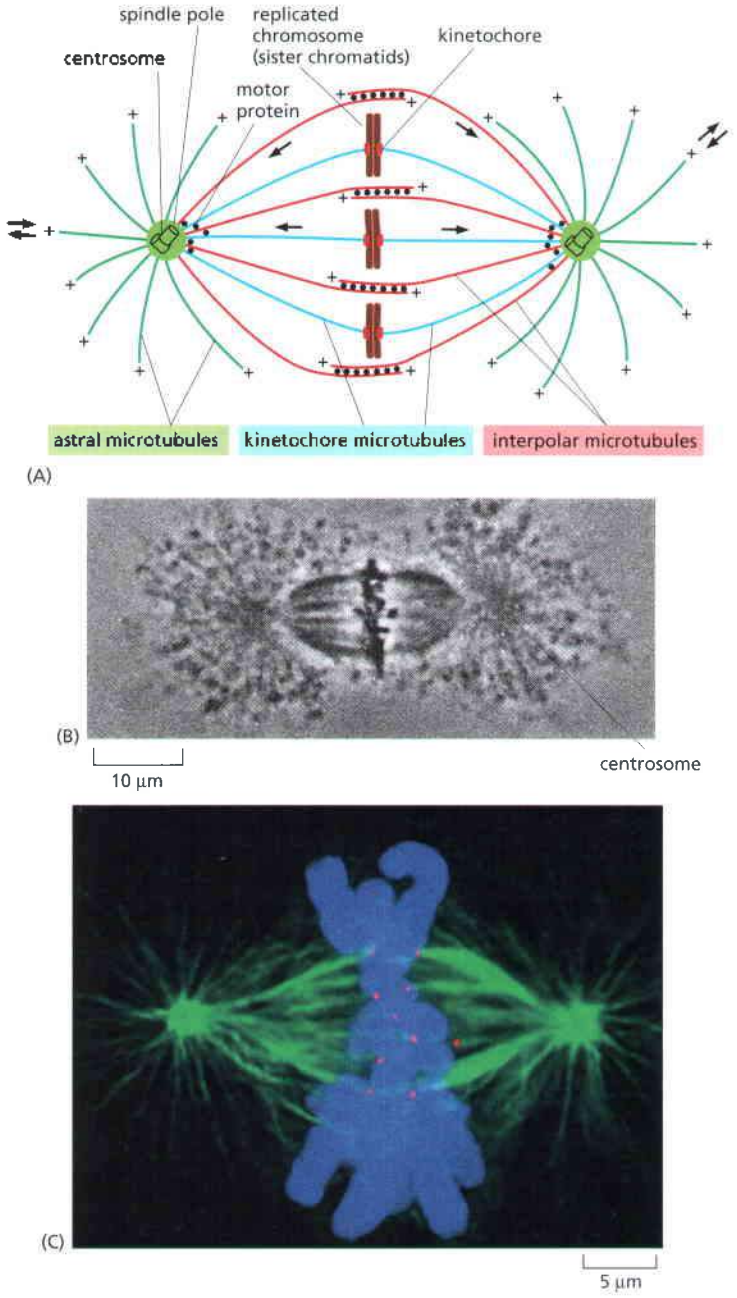


Figure 16–85 The mitotic spindle in animal cells. <GTCT> (A) There are three classes of dynamic microtubules in the mitotic spindle at metaphase: kinetochore microtubules (*blue*) that attach each chromosome to the spindle pole, interpolar microtubules (*red*) that hold the two halves of the spindle together, and astral microtubules (*green*) that can interact with the cell cortex. All of the microtubules are oriented with their minus ends at the spindle poles where the centrosomes reside, and their plus ends projecting away. As indicated by the arrows, the astral microtubules undergo dynamic instability, growing and shrinking at their plus ends, while the kinetochore microtubules and interpolar microtubules both undergo continuous flux toward the spindle poles. (B) A phase-contrast micrograph of an isolated mitotic spindle at metaphase, with the chromosomes aligned at the spindle equator. (C) This fluorescence micrograph shows the microtubules of the spindle in *green* and the chromosomes in *blue*. The *red* spots mark the kinetochores, specialized structures that connect the microtubules to the chromosomes. (B, from E.D. Salmon and R.R. Segall, *J. Cell Biol.* 86:355–365, 1980. With permission from The Rockefeller University Press; C, from A. Desai, *Curr. Biol.* 10:R508, 2000. With permission from Elsevier.)

After the bipolar mitotic spindle has assembled, it can appear stable and quiescent for long periods of time. In many animals, the unfertilized egg arrests its cell cycle in meiotic metaphase, and the spindle waits for days or months until fertilization triggers the progression of the cell cycle (see Chapter 21). This steady appearance is deceptive, because the spindle is actually an extremely dynamic structure, tensed for action that will begin when the chromosomes suddenly begin to separate in anaphase. For example, many of the spindle microtubules exhibit a behavior called *poleward flux*, with a net addition of tubulin subunits at their plus ends, balancing a net loss at their minus ends near the spindle pole. Poleward flux is driven by the action of minus end-directed motor proteins at the spindle pole that are constantly reeling in the microtubules, and the bipolar plus end-directed kinesin-5 motors on the interpolar microtubules that are constantly pushing them apart (see Figure 16–85). As will be discussed in Chapter 17, the delicate balance between these two types of motor protein activities in the spindle also determines its length. Overall the mitotic spindle represents a collaborative effort combining the dynamic properties of microtubules with the individual actions of dozens of molecular motors and other organizing components.

Many Cells Can Crawl Across a Solid Substratum

The process of cell crawling provides another instance where we can appreciate the dynamic integration of cytoskeletal filaments, filament regulators, and motor proteins. Many cells move by crawling over surfaces rather than by using cilia or flagella to swim. Predatory amoebae crawl continuously in search of food, and they can easily be observed to attack and devour smaller ciliates and flagellates in a drop of pond water. In animals, almost all cell locomotion occurs by crawling, with the notable exception of swimming sperm. During embryogenesis, the structure of an animal is created by the migrations of individual cells to specific target locations and by the coordinated movements of whole epithelial sheets (discussed in Chapter 23). In vertebrates, *neural crest cells* are remarkable for their long-distance migrations from their site of origin in the neural tube to a variety of sites throughout the embryo. These cells have diverse fates, becoming skin pigment cells, sensory and sympathetic neurons and glia, and various structures of the face. Long-distance crawling is fundamental to the construction of the entire nervous system: it is in this way that the actin-rich growth cones at the advancing tips of developing axons travel to their eventual synaptic targets, guided by combinations of soluble signals and signals bound to cell surfaces and extracellular matrix along the way.

The adult animal also seethes with crawling cells. Macrophages and neutrophils crawl to sites of infection and engulf foreign invaders as a critical part of the innate immune response. Osteoclasts tunnel into bone, forming channels that are filled in by the osteoblasts that follow after them, in a continuous process of bone remodeling and renewal. Similarly, fibroblasts can migrate through connective tissues, remodeling them where necessary and helping to rebuild damaged structures at sites of injury. In an ordered procession, the cells in the epithelial lining of the intestine travel up the sides of the intestinal villi, replacing absorptive cells lost at the tip of the villus. Unfortunately, cell crawling also has a role in many cancers, when cells in a primary tumor invade neighboring tissues and crawl into blood vessels or lymph vessels and then emerge at other sites in the body to form metastases.

Cell crawling is a highly complex integrated process, dependent on the actin-rich cortex beneath the plasma membrane. Three distinct activities are involved: *protrusion*, in which actin-rich structures are pushed out at the front of the cell; *attachment*, in which the actin cytoskeleton connects across the plasma membrane to the substratum; and *traction*, in which the bulk of the trailing cytoplasm is drawn forward (**Figure 16–86**). In some crawling cells, such as

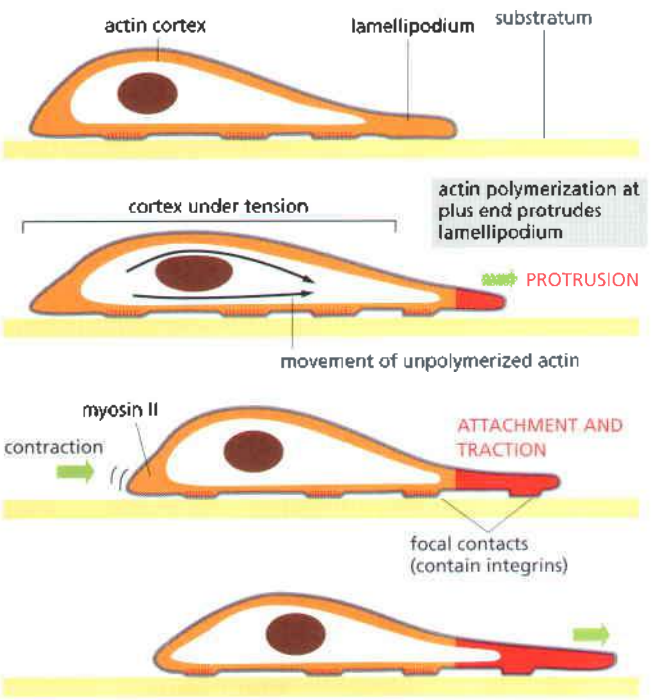


Figure 16–86 A model of how forces generated in the actin-rich cortex move a cell forward. The actin-polymerization-dependent protrusion and firm attachment of a lamellipodium at the leading edge of the cell moves the edge forward (*green* arrows at front) and stretches the actin cortex. Contraction at the rear of the cell propels the body of the cell forward (*green* arrow at back) to relax some of the tension (traction). New focal contacts are made at the front, and old ones are disassembled at the back as the cell crawls forward. The same cycle can be repeated, moving the cell forward in a stepwise fashion. Alternatively, all steps can be tightly coordinated, moving the cell forward smoothly. The newly polymerized cortical actin is shown in *red*.

keratocytes from the fish epidermis, these activities are closely coordinated, and the cells seem to glide forward smoothly without changing shape. In other cells, such as fibroblasts, these activities are more independent, and the locomotion is jerky and irregular.

Actin Polymerization Drives Plasma Membrane Protrusion

The first step in locomotion, protrusion of a leading edge, seems to rely primarily on forces generated by actin polymerization pushing the plasma membrane outward. Different cell types generate different types of protrusive structures, including filopodia (also known as microspikes), lamellipodia, and pseudopodia. All are filled with a dense core of filamentous actin, which excludes membrane-enclosed organelles. The three structures differ primarily in the way in which the actin is organized—in one, two, or three dimensions, respectively—and we have already discussed how this results from the presence of different actin-associated proteins.

Filopodia, formed by migrating growth cones and some types of fibroblasts, are essentially one-dimensional. They contain a core of long, bundled actin filaments, which are reminiscent of those in microvilli but longer and thinner, as well as more dynamic. **Lamellipodia**, formed by epithelial cells and fibroblasts, as well as by some neurons, are two-dimensional, sheet-like structures. They contain an orthogonally cross-linked mesh of actin filaments, most of which lie in a plane parallel to the solid substratum. **Pseudopodia**, formed by amoebae and neutrophils, are stubby three-dimensional projections filled with an actin-filament gel. <ATGG> Perhaps because their two-dimensional geometry is most convenient for examination with the light microscope, we have more information about the dynamic organization and protrusion mechanism of lamellipodia than we have for either filopodia or pseudopodia.

Lamellipodia contain all of the machinery that is required for cell motility. They have been especially well studied in the epithelial cells of the epidermis of fish and frogs, which are known as *keratocytes* because of their abundant keratin filaments. These cells normally cover the animal by forming an epithelial sheet, and they are specialized to close wounds very rapidly, moving at rates up to 30 μm/min. When cultured as individual cells, keratocytes assume a distinctive shape with a very large lamellipodium and a small, trailing cell body that is not attached to the substratum (**Figure 16–87**). Fragments of this lamellipodium can be sliced off with a micropipette. Although the fragments generally lack microtubules and membrane-enclosed organelles, they continue to crawl normally, looking like tiny keratocytes.

The dynamic behavior of actin filaments can be studied in keratocyte lamellipodia by marking a small patch of actin and examining its fate. This reveals

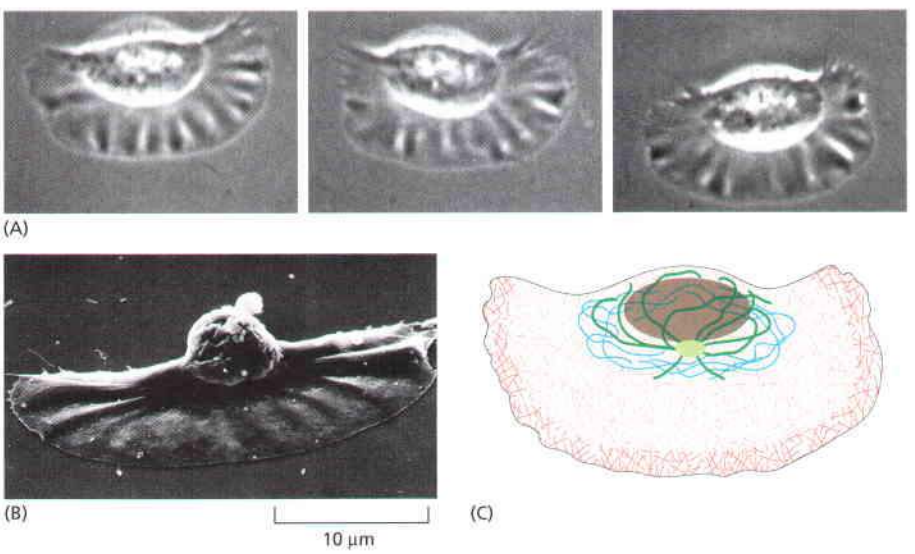


Figure 16–87 Migratory keratocytes from a fish epidermis. <GTTA> (A) Light micrographs of a keratocyte in culture, taken about 15 sec apart. This cell is moving at about 15 μm/sec. (B) Keratocyte seen by scanning electron microscopy, showing its broad, flat lamellipodium and small cell body, including the nucleus, carried up above the substratum at the rear. (C) Distribution of cytoskeletal filaments in this cell. Actin filaments *(red)* fill the large lamellipodium and are responsible for the cell's rapid movement. Microtubules *(green)* and intermediate filaments *(blue)* are restricted to the regions close to the nucleus. (A and B, courtesy of Juliet Lee.)

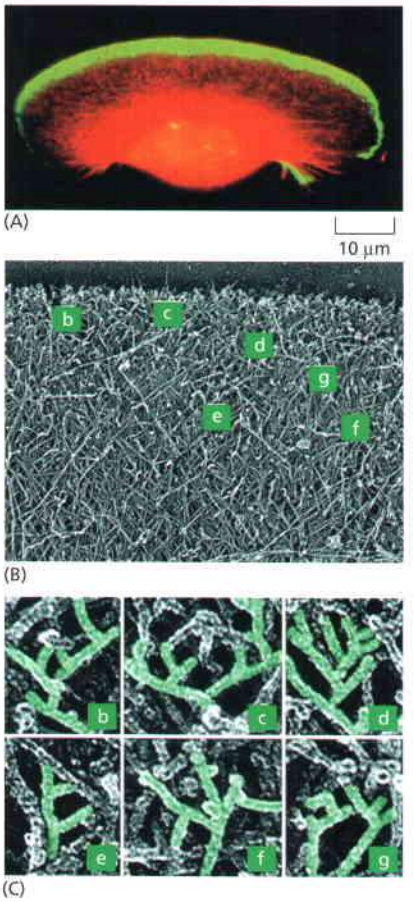


Figure 16–88 Actin filament nucleation and web formation by the ARP complex in lamellipodia. (A) A keratocyte with actin filaments labeled in *red* by fluorescent phalloidin, and the ARP complex labeled in *green* with an antibody raised against one of its component proteins. The regions where the two overlap appear *yellow*. The ARP complex is highly concentrated near the front of the lamellipodium, where actin nucleation is most active. (B) Electron micrograph of a platinum-shadowed replica of the leading edge of a keratocyte, showing the dense actin filament meshwork. The labels denote areas enlarged in C. (C) Close-up views of the marked regions of the actin web at the leading edge shown in B. Numerous branched filaments can be seen, with the characteristic 70° angle formed when the ARP complex nucleates a new actin filament off the side of a preexisting filament (see Figure 16–34). (From T. Svitkina and G. Borisy, *J. Cell Biol.* 145:1009–1026, 1999. With permission from The Rockefeller University Press.)

that, while the lamellipodia crawl forward, the actin filaments remain stationary with respect to the substrate. The actin filaments in the meshwork are mostly oriented with their plus ends facing forward. The minus ends are frequently attached to the sides of other actin filaments by ARP complexes (see Figure 16–34), helping to form the two-dimensional web (**Figure 16–88**). The web as a whole seems to be undergoing treadmilling, assembling at the front and disassembling at the back, reminiscent of the treadmilling that occurs in individual actin filaments and microtubules discussed previously (see Figure 16–14). Treadmilling of a dendritic web built by the ARP complex is only one of several ways that cells can use dynamic actin filaments to drive the protrusion of the leading edge. Some slowly-moving cells including fibroblasts appear to use a mechanism that does not depend on the ARP complex, but still requires coordinated actin filament assembly and disassembly, possibly coordinated by formins.

Maintenance of unidirectional motion by lamellipodia is thought to require the cooperation and mechanical integration of several factors. Filament nucleation is localized at the leading edge, with new actin filament growth occurring primarily in that location to push the plasma membrane forward. Most filament depolymerization occurs at sites located well behind the leading edge. Because *cofilin* (see Figure 16–42) binds cooperatively and preferentially to actin filaments containing ADP-actin (the D form), the new T-form filaments generated at the leading edge should be resistant to depolymerization by cofilin (**Figure 16–89**). As the filaments age and ATP hydrolysis proceeds, cofilin can efficiently disassemble the older filaments. Thus, the delayed ATP hydrolysis by filamentous actin is thought to provide the basis for a mechanism that maintains an efficient, unidirectional treadmilling process in the lamellipodium (**Figure 16–90**). Finally, bipolar myosin II filaments seem to associate with the actin filaments in the web and pull them into a new orientation—from nearly perpendicular to the leading edge to an orientation almost parallel to the leading edge. This contraction prevents protrusion and it pinches in the sides of the locomoting lamellipodium, helping to gather in the sides of the cell as it moves forward (**Figure 16–91**).

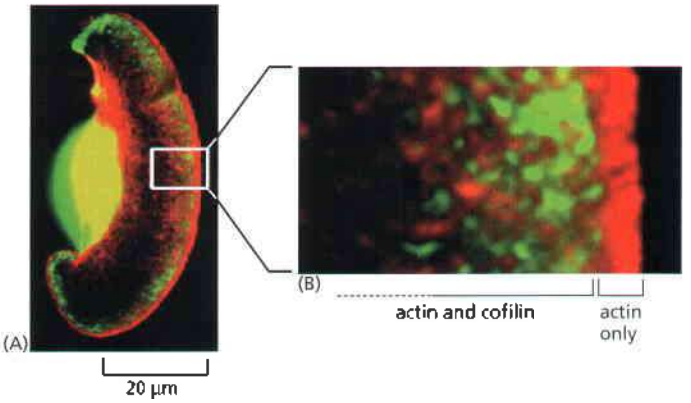


Figure 16–89 Cofilin in lamellipodia. (A) A keratocyte with actin filaments labeled in *red* by fluorescent phalloidin and cofilin labeled in *green* with a fluorescent antibody. The regions where the two overlap appear *yellow*. Although the dense actin meshwork reaches all the way through the lamellipodium, cofilin is not found at the very leading edge. (B) Close-up view of the region marked with the *white* rectangle in A. The actin filaments closest to the leading edge, which are also the ones that have formed most recently and that are most likely to contain ATP actin (rather than ADP actin) in the filament lattice are generally not associated with cofilin. (From T. Svitkina and G. Borisy, *J. Cell Biol.* 145:1009–1026, 1999. With permission from The Rockefeller University Press.)

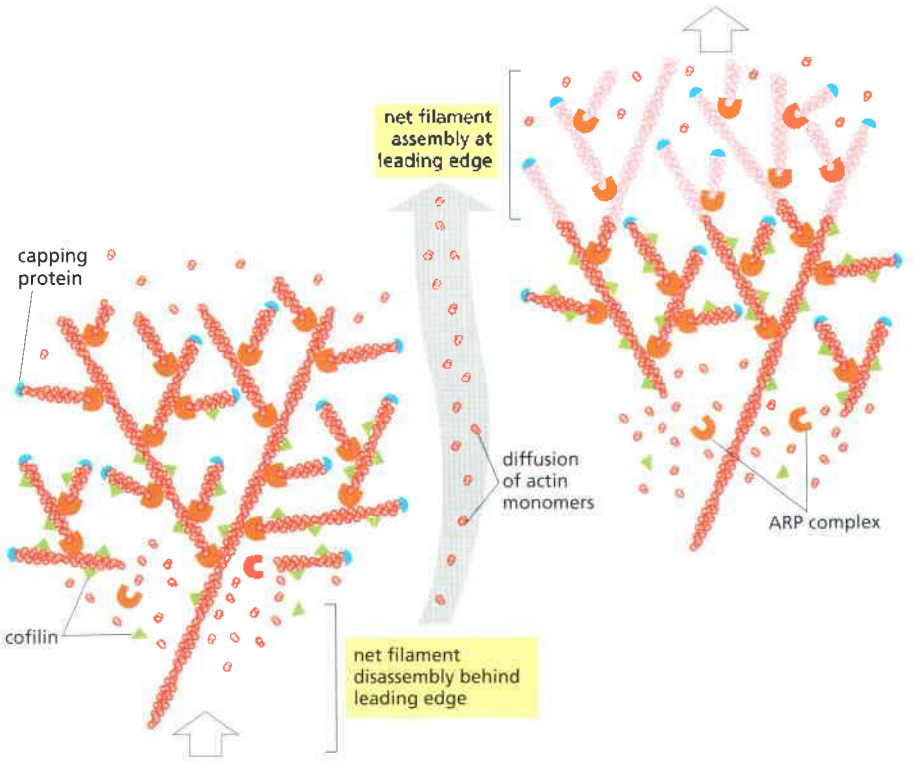
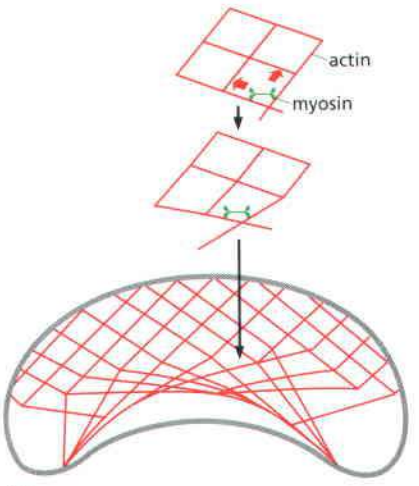


Figure 16–90 A model for protrusion of the actin meshwork at the leading edge. Two time points during advance of the lamellipodium are illustrated, with newly assembled structures at the later time point shown in a lighter color. Nucleation is mediated by the ARP complex at the front. Newly nucleated actin filaments are attached to the sides of preexisting filaments, primarily at a 70° angle. Filaments elongate, pushing the plasma membrane forward because of some sort of anchorage of the array behind. At a steady rate, actin filament plus ends become capped. After newly polymerized actin subunits hydrolyze their bound ATP in the filament lattice, the filaments become susceptible to depolymerization by cofilin. This cycle causes a spatial separation between net filament assembly at the front and net filament disassembly at the rear, so that the actin filament network as a whole can move forward, even though the individual filaments within it remain stationary with respect to the substratum.

The pushing force created by the polymerization of a branched web of actin filaments plays an important role in many cell processes. The polymerization at the plus end can push the plasma membrane outward, as in the example just discussed (see Figure 16–90), or it can propel vesicles or particles through the cell cytoplasm, as in the example of the bacterium *Listeria monocytogenes* discussed in Chapter 24 (see Figure 24–37 <GTAT>). Moreover, when anchored in a more complex way to the membrane, the same type of force drives plasma membrane invaginations, as it does during the endocytotic and phagocytotic processes discussed in Chapter 13.

It is interesting to compare the organization of the actin-rich lamellipodium to the organization of the microtubule-rich mitotic spindle. In both cases, the cell harnesses and amplifies the intrinsic dynamic behavior of the cytoskeletal filament systems to generate large-scale structures that determine the behavior of the whole cell. Both structures feature rapid turnover of their constituent cytoskeletal filaments, even though the structures themselves may remain intact at steady state for long periods of time. The leading edge plasma membrane in the lamellipodium fulfills an organizational role analogous to the condensed chromosomes in organizing and stimulating the dynamics of the mitotic spindle. In both cases, molecular motor proteins help to enhance cytoskeletal filament flux and turnover in the large-scale arrays.

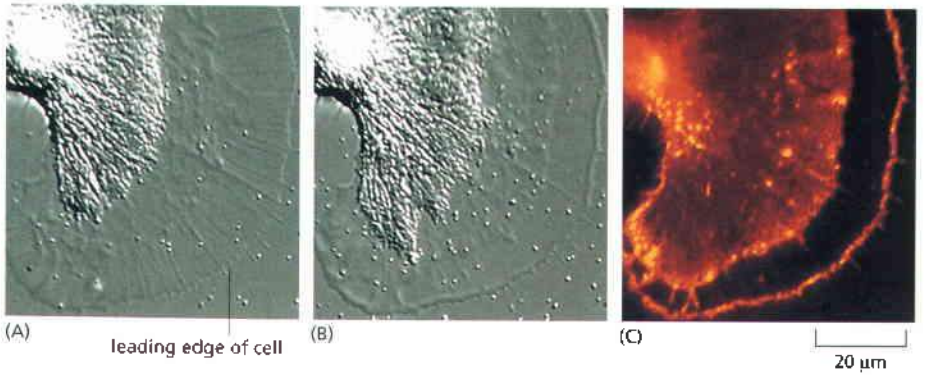


(A)



(B)

Figure 16–91 Contribution of myosin II to polarized cell motility. (A) Myosin II bipolar filaments bind to actin filaments in the dendritic lamellipodial meshwork and cause network contraction. The myosin-driven reorientation of the actin filaments in the dendritic meshwork forms an actin bundle that recruits more myosin II and contributes to generating the contractile forces required for retraction of the trailing edge of the moving cell. (B) A fragment of the large lamellipodium of a keratocyte can be separated from the main cell body either by surgery with a micropipette or by treating the cell with certain drugs. Many of these fragments continue to move rapidly, with the same overall cytoskeletal organization as the intact keratocytes. Actin *(blue)* forms a protrusive meshwork at the front of the fragment. Myosin II *(pink)* is gathered into a band at the rear. (From A. Verkovsky et al., *Curr. Biol.* 9:11–20, 1999. With permission from Elsevier.)



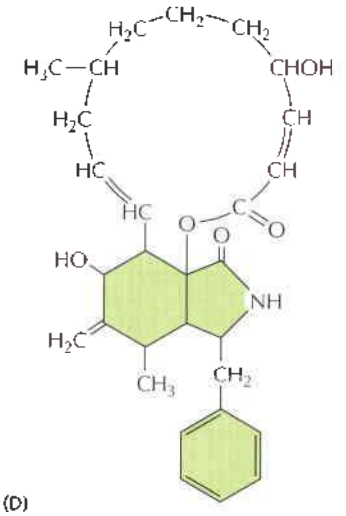


Figure 16–92 **Rearward movement of the actin network in a growth-cone lamellipodium.** (A) A growth cone from a neuron of the sea slug *Aplysia* is cultured on a highly adhesive substratum and viewed by differential-interference-contrast microscopy. Microtubules and membrane-enclosed organelles are confined to the bright, rear area of the growth cone (to the *left*), while a meshwork of actin filaments fills the lamellipodium (on the *right*). (B) After brief treatment with the drug cytochalasin, which caps the plus ends of actin filaments (see Table 16–2, p. 988), the actin meshwork has detached from the front edge of the lamellipodium and has been pulled backward. (C) At the time point shown in B, the cell was fixed and labeled with fluorescent phalloidin to show the distribution of the actin filaments. Some actin filaments persist at the leading edge, but the region behind the leading edge is devoid of filaments. Note the sharp boundary of the rearward-moving actin meshwork. (D) The complex cyclic structure of cytochalasin B. (A–C, courtesy of Paul Forscher.)

Cell Adhesion and Traction Allow Cells to Pull Themselves Forward

Lamellipodia of all cells seem to share a basic, simple type of dynamic organization where actin filament assembly occurs preferentially at the leading edge and actin filament disassembly occurs preferentially at the rear. However, the interactions between the cell and its normal physical environment usually make the situation considerably more complex than for fish keratocytes crawling on a culture dish. Particularly important in locomotion is the intimate crosstalk between the cytoskeleton and the cell adhesion apparatus. Although some degree of adhesion to the substratum is necessary for any form of cell crawling, adhesion and locomotion rate seem generally to be inversely related, with highly adhesive cells moving more slowly than weakly adhesive ones. Keratocytes are so weakly adhesive to the substratum that the force of actin polymerization can push the leading edge forward very rapidly. In contrast, neurons from the sea slug *Aplysia* cultured on a sticky substratum form large lamellipodia that become stuck too tightly to move forward. In these lamellipodia, the same cycle of localized nucleation of new actin filaments, depolymerization of old filaments, and myosin-dependent contraction continues to operate. But because the leading edge is prevented physically from moving forward, the entire actin mesh moves backward toward the cell body instead, pulled by myosins (**Figure 16–92**). The adhesion of most cells lies somewhere between these two extremes, and most lamellipodia exhibit some combination of forward actin filament protrusion (like keratocytes) and rearward actin flux (like the *Aplysia* neurons).

As a lamellipodium, filopodium, or pseudopodium extends forward over a substratum, it can form new attachment sites at the cell front that remain stationary as the cell moves forward over them, persisting until the rear of the cell catches up with them. When an individual lamellipodium fails to adhere to the

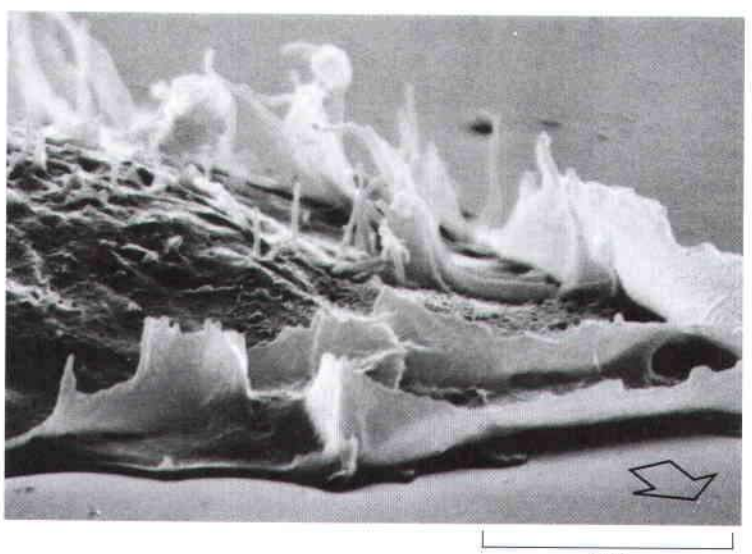


Figure 16–93 **Lamellipodia and ruffles at the leading edge of a human fibroblast migrating in culture.** The *arrow* in this scanning electron micrograph shows the direction of cell movement. As the cell moves forward, lamellipodia that fail to attach to the substratum are swept backward over the dorsal surface of the cell, a movement known as ruffling. (Courtesy of Julian Heath.)

substratum, it is usually lifted up onto the dorsal surface of the cell and rapidly carried backward as a "ruffle" (**Figure 16–93**).

The attachment sites established at the leading edge serve as anchorage points, which allow the cell to generate traction on the substratum and pull its body forward. Myosin motor proteins, especially myosin II, seem to generate traction forces. In many locomoting cells, myosin II is highly concentrated at the posterior of the cell where it may help to push the cell body forward like toothpaste being squeezed out of a tube from the rear (**Figure 16–94**; see also Figure 16–91). *Dictyostelium* amoebae that are deficient in myosin II are able to protrude pseudopodia at normal speeds, but the translocation of their cell body is much slower than that of wild-type amoebae, indicating the importance of myosin II contraction in this part of the cell locomotion cycle. In addition to helping to push the cell body forward, contraction of the actin-rich cortex at the rear of the cell may selectively weaken the older adhesive interactions that tend to hold the cell back. Myosin II may also transport cell body components forward over a polarized array of actin filaments.

The traction forces generated by locomoting cells exert a significant pull on the substratum (**Figure 16–95**). In a living animal, most crawling cells move across a semiflexible substratum made of extracellular matrix, which can be deformed and rearranged by these cell forces. In culture, movement of fibroblasts through a gel of collagen fibrils aligns the collagen, generating an organized extracellular matrix that in turn affects the shape and direction of locomotion of the fibroblasts within it (**Figure 16–96**). Conversely, mechanical tension or stretching applied externally to a cell will cause it to assemble stress fibers and focal adhesions, and become more contractile. Although poorly understood, this two-way mechanical interaction between cells and their physical environment is thought to be a primary way that vertebrate tissues organize themselves.



Figure 16–94 The localization of myosin I and myosin II in a normal crawling *Dictyostelium* amoeba. This cell was crawling toward the upper right at the time that it was fixed and labeled with antibodies specific for two myosin isoforms. Myosin I *(green)* is mainly restricted to the leading edge of pseudopodia at the front of the cell. Myosin II *(red)* is highest in the posterior, actin-rich cortex. Contraction of the cortex at the posterior of the cell by myosin II may help to push the cell body forward. (Courtesy of Yoshio Fukui.)

Members of the Rho Protein Family Cause Major Rearrangements of the Actin Cytoskeleton

Cell migration is one example of a process that requires long-distance communication and coordination between one end of a cell and the other. During directed migration, it is important that the front end of the cell remain structurally and functionally distinct from the back end. In addition to driving local mechanical processes such as protrusion at the front and retraction at the rear, the cytoskeleton is responsible for coordinating cell shape, organization, and mechanical properties from one end of the cell to the other, a distance which is typically several tens of micrometers for animal cells. In many cases, including but not limited to cell migration, large-scale cytoskeletal coordination takes the form of the establishment of cell polarity, where a cell builds different structures with distinct molecular components at the front vs. the back, or at the top vs. the

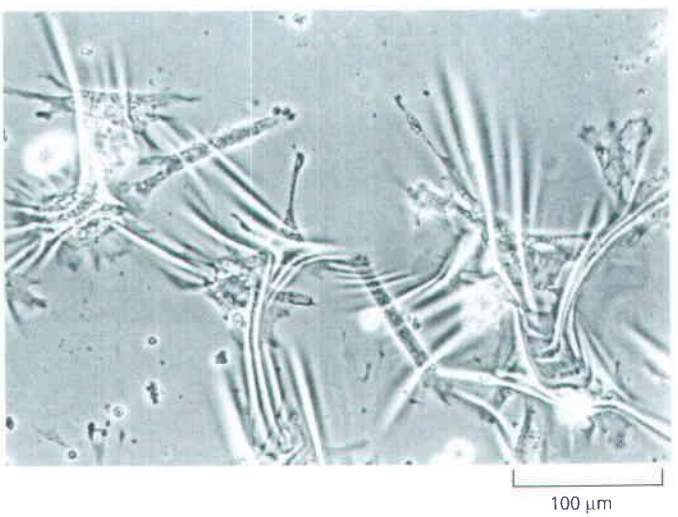


Figure 16–95 Adhesive cells exert traction forces on the substratum. These fibroblasts have been cultured on a very thin sheet of silicon rubber. Attachment of the cells, followed by contraction of their cytoskeleton, has caused the rubber substratum to wrinkle. (From A.K. Harris, P. Wild and D. Stopak, *Science* 208:177–179, 1980. With permission from AAAS.)

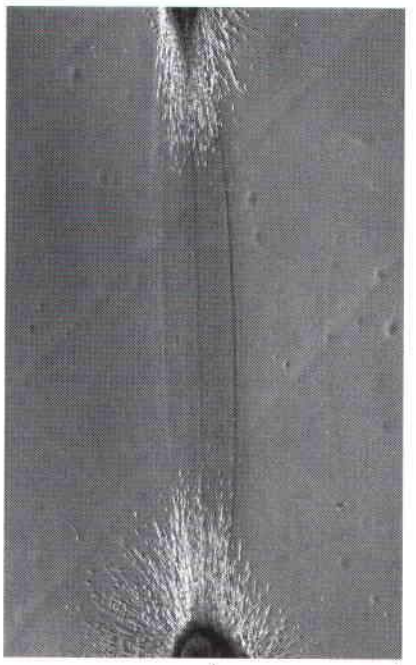


Figure 16–96 Shaping of the extracellular matrix by cell pulling. This micrograph shows a region between two pieces of embryonic chick heart (tissue explants rich in fibroblasts and heart muscle cells) that were grown in culture on a collagen gel for 4 days. A dense tract of aligned collagen fibers has formed between the two explants, apparently as a result of fibroblasts tugging on the collagen. (From D. Stopak and A.K. Harris, *Dev. Biol.* 90:383–398, 1982. With permission from Academic Press.)

bottom. Cell locomotion requires an initial polarization of the cell to set it off in a particular direction. Carefully controlled cell polarization processes are also required for oriented cell divisions in tissues and for formation of a coherent, organized multicellular structure. Genetic studies in yeast, flies, and worms have provided most of our current understanding of the molecular basis of cell polarity. The mechanisms that generate cell polarity in vertebrates are only beginning to be explored. In all known cases, however, the cytoskeleton has a central role, and many of the molecular components have been evolutionarily conserved.

For the actin cytoskeleton, diverse cell-surface receptors trigger global structural rearrangements in response to external signals. But all of these signals seem to converge inside the cell on a group of closely related monomeric GTPases that are members of the **Rho protein family**—*Cdc42*, *Rac*, and *Rho*. The same Rho family proteins are also involved in the establishment of many kinds of cell polarity.

Like other members of the Ras superfamily, these Rho proteins act as molecular switches to control cell processes by cycling between an active, GTP-bound state and an inactive, GDP-bound state (see Figure 3–71). Activation of Cdc42 on the plasma membrane triggers actin polymerization and bundling to form either filopodia or shorter cell protrusions called microspikes. Activation of Rac promotes actin polymerization at the cell periphery leading to the formation of sheet-like lamellipodial extensions and membrane ruffles, which are actin-rich protrusions on the cell's dorsal surface (see Figure 16–93). Activation of Rho promotes both the bundling of actin filaments with myosin II filaments into stress fibers and the clustering of integrins and associated proteins to form focal contacts (**Figure 16–97**). These dramatic and complex structural changes occur because each of these three molecular switches has numerous downstream target proteins that affect actin organization and dynamics.

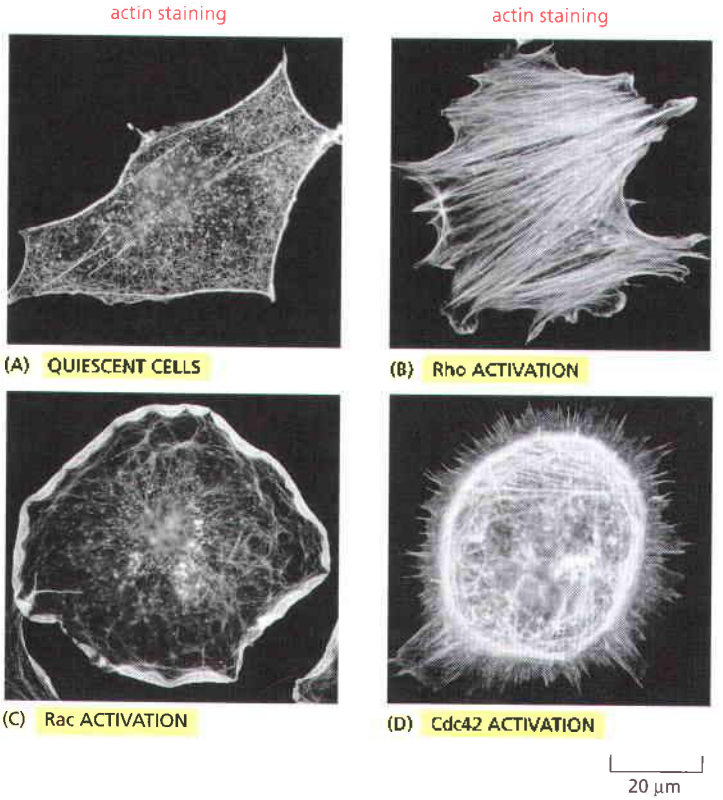


Figure 16–97 The dramatic effects of Rac, Rho, and Cdc42 on actin organization in fibroblasts. In each case, the actin filaments have been labeled with fluorescent phalloidin. (A) Serum-starved fibroblasts have actin filaments primarily in the cortex, and relatively few stress fibers. (B) Microinjection of a constitutively activated form of Rho causes the rapid assembly of many prominent stress fibers. (C) Microinjection of a constitutively activated form of Rac, a closely related monomeric GTPase, causes the formation of an enormous lamellipodium that extends from the entire circumference of the cell. (D) Microinjection of a constitutively activated form of Cdc42, another Rho family member, causes the protrusion of many long filopodia at the cell periphery. The distinct global effects of these three GTPases on the organization of the actin cytoskeleton are mediated by the actions of dozens of other protein molecules that are regulated by the GTPases. These target proteins include some of the various actin-associated proteins that we have discussed in this chapter. (From A. Hall, *Science* 279:509–514, 1998. With permission from AAAS.)

Some key targets of activated Cdc42 are members of the WASp protein family. Human patients deficient in WASp suffer from Wiskott-Aldrich Syndrome, a severe form of immunodeficiency where immune system cells have abnormal actin-based motility and platelets do not form normally. Although WASp itself is expressed only in blood cells and immune system cells, other family members are expressed ubiquitously that enable activated Cdc42 to enhance actin polymerization. **WASp proteins** can exist in an inactive folded conformation and an activated open conformation. Association with Cdc42-GTP stabilizes the open form of WASp, enabling it to bind to the ARP complex and strongly enhancing this complex's actin-nucleating activity (see Figure 16–34). In this way, activation of Cdc42 increases actin nucleation.

Rac-GTP also activates WASp family members, as well as activating the crosslinking activity of the gel-forming protein filamin, and inhibiting the contractile activity of the motor protein myosin II, stabilizing the lamellipodia and inhibiting the formation of contractile stress fibers (**Figure 16–98**A).

Rho-GTP has a very different set of targets. Instead of activating the ARP complex to build actin networks, Rho-GTP turns on formin proteins to construct parallel actin bundles. At the same time, Rho-GTP activates a protein kinase that indirectly inhibits the activity of cofilin, leading to actin filament stabilization. The same protein kinase inhibits a phosphatase acting on myosin light chains (see Figure 16–72). The consequent increase in the net amount of myosin light chain phosphorylation increases the amount of contractile myosin motor protein activity in the cell, enhancing the formation of tension-dependent structures such as stress fibers (Figure 16–98B).

In some cell types, Rac-GTP activates Rho, usually with kinetics that are slow compared to Rac's activation of the ARP complex. This enables cells to use the Rac pathway to build a new actin structure while subsequently activating the Rho pathway to induce a contractility that builds up tension in this structure. This occurs, for example, during the formation and maturation of cell-cell contacts. As we will explore in more detail below, the communication between the Rac and Rho pathways also facilitates maintenance of the large-scale differences between the cell front and the cell rear during migration.

Extracellular Signals Can Activate the Three Rho Protein Family Members

The activation of the monomeric GTPases Rho, Rac, and Cdc42 occurs through an exchange of GTP for a tightly bound GDP molecule, catalyzed by guanine nucleotide exchange factors (GEFs). Of the 85 GEFs that have been identified in

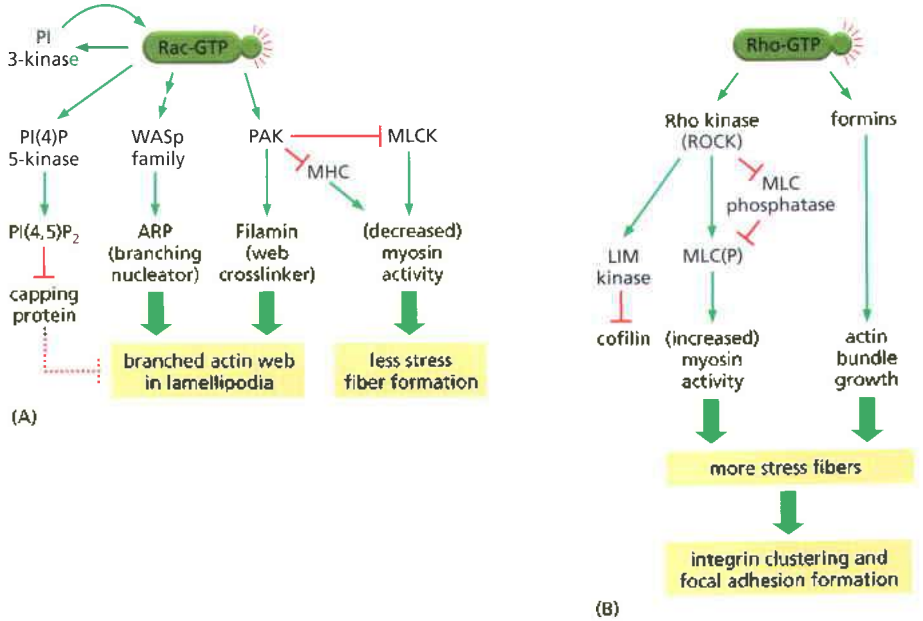


Figure 16–98 The contrasting effects of Rac and Rho activation on actin organization. (A) Activation of the small GTPase Rac leads to actin nucleation by the ARP complex and other alterations in actin accessory proteins that tend to favor the formation of actin networks, as in lamellipodia. Several different pathways contribute independently. Rac-GTP activates members of the WASp protein family, which in turn activate actin nucleation and branched web formation by the ARP complex. In a parallel pathway, Rac-GTP activates a protein kinase, PAK, which has several targets including the web-forming crosslinker filamin, which is activated by phosphorylation, and the myosin light chain kinase (MLCK), which is inhibited by phosphorylation. The resulting decrease in phosphorylation of the myosin regulatory light chain leads to myosin II filament disassembly and a decrease in contractile activity. In some cells, PAK also directly inhibits myosin II activity by phosphorylation of the myosin heavy chain (MHC). Another set of pathways downstream of Rac activation is mediated by phosphoinositide lipid signals. Local creation of PIP_2 [$PI(4,5)P_2$] may help to reduce the activity of capping protein, to further aid actin polymerization. Activation of PI 3-kinase, which generates PIP_3 from PIP_2, leads to further activation of Rac itself via a positive feedback loop.

(B) Activation of the related GTPase Rho leads to nucleation of actin filaments by formins and increases contraction by myosin II, promoting the formation of contractile actin bundles such as stress fibers. Activation of myosin II by Rho requires a Rho-dependent protein kinase called Rock. This kinase inhibits the phosphatase that removes the activating phosphate groups from myosin II light chains (MLC); it may also directly phosphorylate the myosin light chains in some cell types. Rock also activates other protein kinases, such as LIM kinase, which in turn contributes to the formation of stable contractile actin filament bundles by inhibiting the actin depolymerizing factor cofilin. A similar signaling pathway is important for forming the contractile ring necessary for cytokinesis (see Figure 17–52).

the human genome, some are specific for an individual Rho family GTPase, whereas others seem to act on all three family members. The number of GEFs exceeds the number of Rho GTPases that they regulate because different GEFs are restricted to specific tissues and even specific subcellular locations, and they are sensitive to distinct kinds of regulatory inputs. Various cell-surface receptors activate GEFs. An example is the Eph receptor tyrosine kinase involved in neurite growth cone guidance, which is discussed in detail in Chapter 15. Interestingly, several of the Rho family GEFs associate with the growing ends of microtubules by binding to one of the +TIPs. This provides a connection between the dynamics of the microtubule cytoskeleton and the large-scale organization of the actin cytoskeleton, which is important for the overall integration of cell shape and movement.

The Rho family GTPases are also primary determinants of cell polarity in budding yeast, where extensive genetic analyses have increased our understanding of the general mechanisms involved. On starvation, yeasts, like many other unicellular organisms, sporulate. But sporulation can occur only in diploid budding yeast cells, whereas budding yeasts mainly proliferate as haploid cells. A starving haploid individual must therefore locate a partner of the opposite mating type, woo it, and mate with it before sporulating. Yeast cells are unable to swim and, instead, reach their mates by polarized growth. The haploid form of budding yeast comes in two mating types, **a** and α, which secrete mating factors known as **a**-factor and α-factor, respectively. These secreted signal molecules act by binding to cell-surface receptors that belong to the G-protein-coupled receptor superfamily (discussed in Chapter 15). One consequence of the binding of α-factor to its receptor is to cause the recipient cell to become polarized, adopting a shape known as a "shmoo" (**Figure 16–99**). In the presence of an α-factor gradient, the **a**-cell shmoo tip is directed toward the highest concentration of the signal molecule, which under normal circumstances would direct it toward an amorous α cell located nearby.

This polarized cell growth requires alignment of the actin cytoskeleton in response to the mating factor signal. When the signal binds to its receptor, the receptor activates Cdc42, which in turn induces assembly of actin filaments at the location closest to the source of the signal. Local activation of Cdc42 is further enhanced by a positive feedback loop, requiring actin-dependent transport of Cdc42 itself as well as its GEF and other signaling components along the newly assembled actin structures toward the site of the signal. Subsequently, actin cables are assembled pointing toward the site of Cdc42 accumulation due to the activation of another Rho family GTPase that in turn stimulates a yeast formin. The actin cables serve as tracks for directed transport and exocytosis of new cell wall material, resulting in the polarized growth of the shmoo tip (**Figure 16–100**).

Haploid budding yeast cells use this same polarization machinery during vegetative growth. To form the bud that will grow out to become a daughter cell, the yeast must direct new plasma membrane and cell wall material primarily to a single site. As with shmoo formation, this requires an initial cytoskeletal polarity, with most actin patches in the growing bud and actin cables oriented along the bud axis. In haploid cells, a new bud site is always constructed immediately adjacent to the previous bud site. In this case, the spatial cues that set up cytoskeletal polarity are intrinsic to the cell, left behind from previous rounds of cell division. Cdc42 is once again involved in transducing the signal from the destined bud site to the cytoskeleton, and most of the proteins involved in the upstream and downstream pathways have been identified through genetic

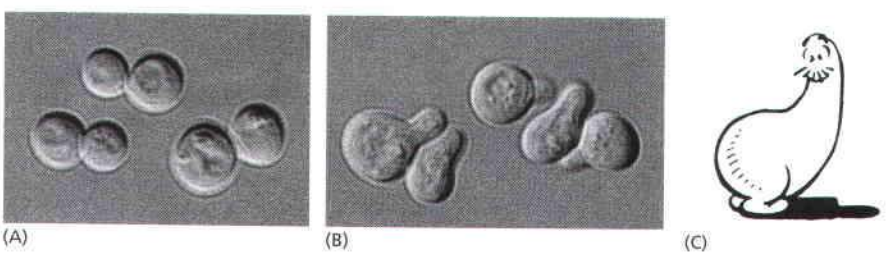


Figure 16–99 Morphological polarization of yeast cells in response to mating factor. (A) Cells of *Saccharomyces cerevisiae* are usually spherical. (B) They become polarized when treated with mating factor from cells of the opposite mating type. The polarized cells are called "shmoos." (C) Al Capp's famous cartoon character, the original Shmoo. (A and B, courtesy of Michael Snyder; C, © 1948 Capp Enterprises, Inc. Used by permission.)

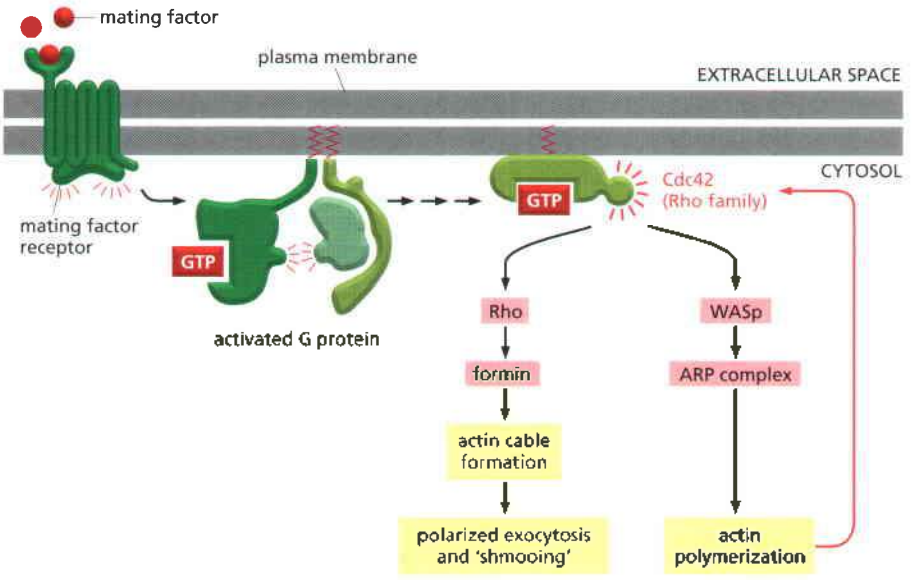


Figure 16–100 The signaling pathway in the yeast mating factor response. The extracellular mating factor binds to a G-protein-coupled receptor in the plasma membrane. Activation of the receptor triggers dissociation of the GTP-bound Gα subunit from a heterotrimeric G-protein (discussed in Chapter 15). This in turn activates the Rho family GTP-binding protein, Cdc42. As in mammalian cells, Cdc42 activates a WASp family protein that activates the ARP complex, leading to local actin nucleation at the site of mating factor binding. The local actin nucleation and filament growth create a positive feedback loop whereby Cdc42 activity is further enhanced. This leads to extensive Rho and formin activation, and finally to actin cable formation, polarized growth, and acquisition of a shmoo shape. In addition, receptor activation triggers other responses through a MAP kinase cascade (discussed in Chapter 15), preparing the haploid cell for mating (not shown).

experiments. Subsequent to their identification in yeast, many of these proteins have been found to have homologs in other organisms, where they are often likewise involved in the establishment of cell polarity.

External Signals Can Dictate the Direction of Cell Migration

Chemotaxis is defined as cell movement in a direction controlled by a gradient of a diffusible chemical. This is a particularly interesting case where external signals trigger the Rho family proteins to set up large-scale cell polarity by influencing the organization of the apparatus required for cell motility, described above. One well-studied example is the chemotactic movement of a class of white blood cells, called *neutrophils*, toward a source of bacterial infection. Receptor proteins on the surface of neutrophils enable them to detect the very low concentrations of the *N*-formylated peptides that are derived from bacterial proteins (only procaryotes begin protein synthesis with *N*-formylmethionine). Using these receptors, neutrophils are guided to bacterial targets by their ability to detect a difference of only 1% in the concentration of these diffusible peptides on one side of the cell versus the other (**Figure 16–101**).

Both in this case and in the similar chemotaxis of *Dictyostelium* amoebae toward a source of cyclic AMP, a local polymerization of actin near the receptors is stimulated when the receptors bind their ligands. This actin polymerization response depends on the monomeric Rho-family GTPases discussed earlier. As in the shmooing yeast (see Figure 16–99), the responding cell extends a protrusion toward the signal. For chemotactic cells, binding of the chemoattractant ligand to its G-protein coupled receptor activates phosphoinositide 3′ kinases (PI3Ks), which generates a lipid-based signaling molecule ($PI(3,4,5)P_3$) that in turn activates the Rac GTPase. Rac then activates the ARP complex and lamellipodial protrusion results (see Figure 16–98). Through an unknown mechanism,

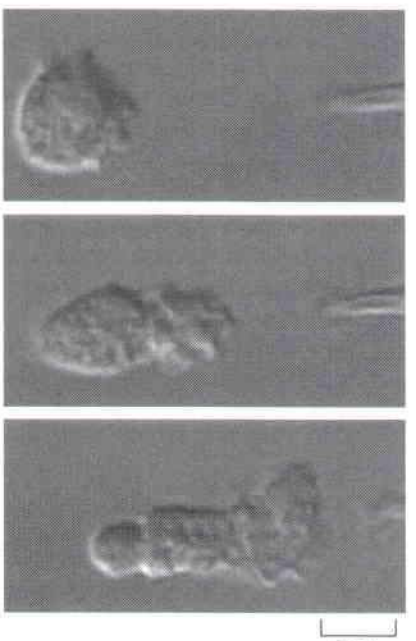


Figure 16–101 Neutrophil polarization and chemotaxis. <GTCG> <TGTA> The pipette tip at the right is leaking a small amount of the peptide formyl-Met-Leu-Phe. Only bacterial proteins have formylated methionine residues, so the human neutrophil recognizes this peptide as the product of a foreign invader (discussed in Chapter 24). The neutrophil quickly extends a new lamellipodium toward the source of the chemoattractant peptide *(top)*. It then extends this lamellipodium and polarizes its cytoskeleton so that contractile myosin II is located primarily at the rear, opposite the position of the lamellipodium *(middle)*. Finally, the cell crawls toward the source of this peptide *(bottom)*. If a real bacterium were the source of the peptide, rather than an investigator's pipette, the neutrophil would engulf the bacterium and destroy it (see also Figure 16–4). (From O.D. Weiner et al., *Nat. Cell Biol.* 1:75–81, 1999. With permission from Macmillan Publishers Ltd.)

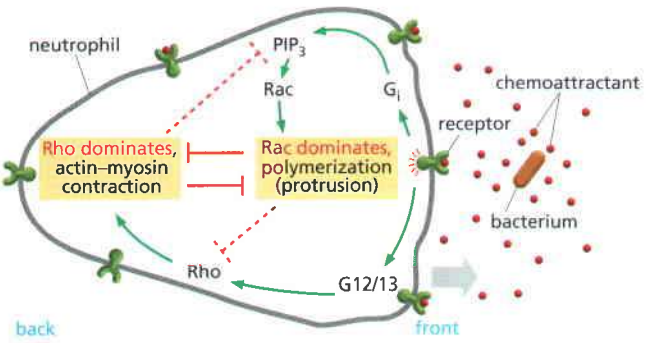


Figure 16–102 Signaling during neutrophil polarization. Bacteria that have invaded the human body secrete molecules that are recognized as foreign by the cells of the immune system, including neutrophils. Binding of the bacterial molecules to G-protein-coupled receptors on the neutrophils stimulate directed motility. These receptors are found all over the surface of the cell, but are more likely to be bound to the bacterial ligand at the front. Two distinct signaling pathways contribute to the cell's polarization. At the front of the cell, close to the source of the bacterial signal, stimulation of the Rac pathway leads, via the trimeric G protein G_i, to growth of protrusive actin networks. Second messengers within this pathway are short-lived, so protrusion is limited to the region of the cell closest to the stimulant. The same receptor also stimulates a second signaling pathway, via the trimeric G proteins G12 and G13 (denoted G12/13), that triggers the activation of Rho. The two pathways are mutually antagonistic. Since Rac-based protrusion is active at the front of the cell, Rho is activated only at the rear of the cell, stimulating contraction of the cell rear and assisting directed movement. For a real-life example of the effectiveness of this signaling system, see Figure 16–4.

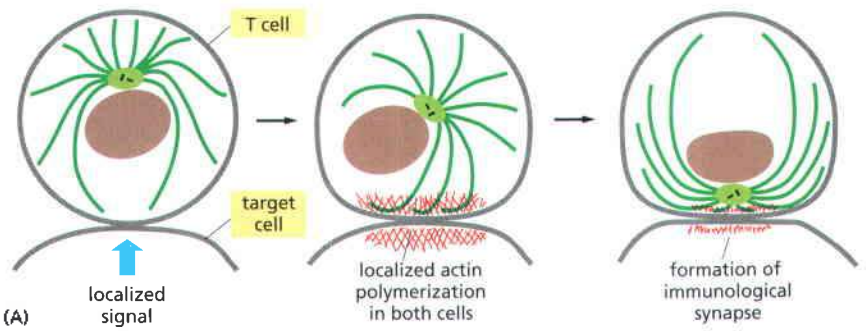
accumulation of the polarized actin web at the leading edge causes further local enhancement of PI3K activity in a positive feedback loop, strengthening the induction of protrusion. The $PI(3,4,5)P_3$ that activates Rac cannot diffuse far from its site of synthesis, since it is rapidly converted back into PIP_2 by a constitutively active lipid phosphatase. At the same time, binding of the chemoattractant ligand to its receptor activates another signaling pathway that turns on Rho and enhances myosin-based contractility. The two processes directly inhibit each other, and as a result, Rac activation dominates in the front of the cell and Rho activation dominates in the rear (**Figure 16–102**). This enables the cell to maintain its functional polarity with protrusion at the leading edge and contraction at the back.

Nondiffusible chemical cues attached to the extracellular matrix or to the surface of cells can also influence the direction of cell migration. When these signals activate receptors, they can cause increased cell adhesion and directed actin polymerization. Most long-distance cell migrations in animals, including neural-crest-cell migration and the travels of neuronal growth cones, depend on a combination of diffusible and non-diffusible signals to steer the locomoting cells or growth cones to their proper destinations (see Figure 15–62).

Communication Between the Microtubule and Actin Cytoskeletons Coordinates Whole-Cell Polarization and Locomotion

To help organize persistent movement in a particular direction cells use their microtubules along with their actin filaments. In many locomoting cells, the position of the centrosome is influenced by the location of protrusive actin polymerization, being found on the forward side of the nucleus. The mechanism of centrosome reorientation is not clear, although there is evidence that the Rho family protein Cdc42 may be involved. It is thought that the activation of receptors on one edge of a cell might not only stimulate actin polymerization there (and therefore local protrusion) but also locally activate dynein-like motor proteins that move the centrosome by pulling on its microtubules. Several effector proteins downstream of Rac and Rho modulate microtubule dynamics directly. For example, a protein kinase activated by Rac can phosphorylate (and therefore inhibit) the tubulin binding protein stathmin (see Panel 16–3, pp. 994–995), destabilizing microtubules, and Rho activation appears to stabilize microtubules.

In turn, microtubule dynamics influence actin rearrangements. The centrosome nucleates a large number of dynamic microtubules, and its repositioning means that many of these microtubules have their plus ends extending from the centrosome into the protrusive region of the cell. The dynamic microtubule plus ends may indirectly modulate local adhesion and also activate the Rac GTPase to further increase actin polymerization in the protrusive region by delivering Rac-GEFs that bind to the +TIPs traveling on growing microtubule ends. The increased concentration of microtubules would thereby encourage further protrusion, creating a positive feedback loop that enables protrusive motility to persist in the same direction for a prolonged period. Regardless of the exact mechanism, the orientation of the centrosome seems to reinforce the polarity information that the actin cytoskeleton receives from the outside world, allowing a sensitive response to weak signals.



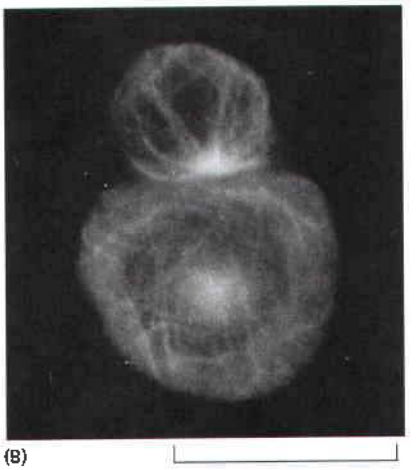


Figure 16–103 The polarization of a cytotoxic T cell after target-cell recognition. (A) Changes in the cytoskeleton of a cytotoxic T cell after it has made contact with a target cell. The initial recognition event results in signals that cause actin polymerization in both cells at the site of contact. In the T cell, interactions between the actin-rich contact zone and microtubules emanating from the centrosome result in reorientation of the centrosome, so that the associated Golgi apparatus is directly apposed to the target cell. (B) Immuno-fluorescence micrograph in which both the T cell *(top)* and its target cell *(bottom)* have been stained with an antibody against microtubules. The centrosome and the microtubules radiating from it in the T cell are oriented toward the point of cell–cell contact. In contrast, the microtubule array in the target cell is not polarized. (B, from B. Geiger, D. Rosen and G. Berke, *J. Cell Biol.* 95:137–143, 1982. With permission from The Rockefeller University Press.)

A similar cooperative feedback loop seems to operate in many other instances of cell polarization. A particularly interesting example is the killing of specific target cells by T lymphocytes. These cells are a critical component of the vertebrate's adaptive immune response to infection by viruses. T cells, like neutrophils, use actin-based motility to crawl through the body's tissue and find infected target cells. When a T cell comes into contact with a virus-infected cell and its receptors recognize foreign viral antigens on the surface of the target cell, the same polarization machinery is engaged in a very different way to facilitate killing of the target cell. Rac is activated at the point of cell–cell contact and causes actin polymerization at this site, creating a specialized region of the cortex. This specialized site causes the centrosome to reorient, moving with its microtubules to the zone of T-cell-target contact (**Figure 16–103**). The microtubules, in turn, position the Golgi apparatus right under the contact zone, focusing the killing machinery onto the target cell. The mechanism of killing is discussed in Chapter 25 (see Figure 25–47).

The Complex Morphological Specialization of Neurons Depends on the Cytoskeleton

For our final case study of the ways that the intrinsic properties of the eucaryotic cytoskeleton enable specific and enormously complicated large-scale cell behaviors, we examine the neuron. Neurons begin life in the embryo as unremarkable cells, which use actin-based motility to migrate to specific locations. Once there, however, they send out a series of long specialized processes that will either receive electrical signals *(dendrites)* or transmit electrical signals *(axons)* to their target cells. <TACC> The beautiful and elaborate branching morphology of axons and dendrites enables neurons to form tremendously complex signaling networks, interacting with many other cells simultaneously and making possible the complicated and often unpredictable behavior of the higher animals. Both axons and dendrites (collectively called *neurites)* are filled with bundles of microtubules that are critical to both their structure and their function.

In axons, all the microtubules are oriented in the same direction, with their minus end pointing back toward the cell body and their plus end pointing forward toward the axon terminals (**Figure 16–104**). The microtubules do not reach

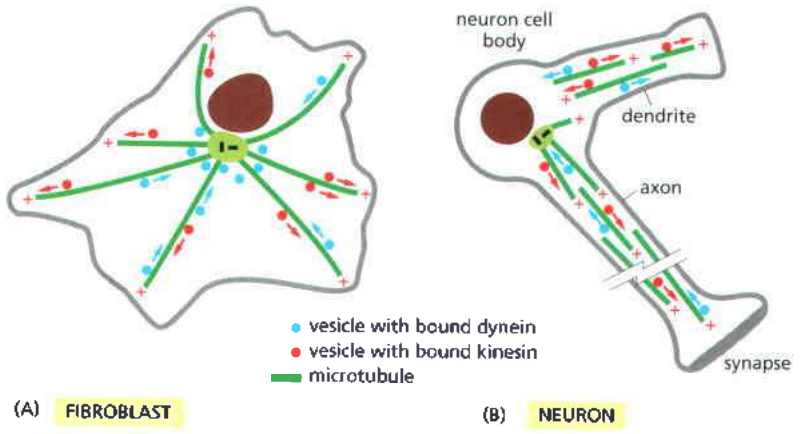


Figure 16–104 Microtubule organization in fibroblasts and neurons. (A) In a fibroblast, microtubules emanate outward from the centrosome in the middle of the cell. Vesicles with plus-end-directed kinesin attached move outward, and vesicles with minus-end-directed dynein attached move inward. (B) In a neuron, microtubule organization is more complex. In the axon, all microtubules share the same polarity, with the plus ends pointing outward toward the axon terminus. No one microtubule stretches the entire length of the axon; instead, short overlapping segments of parallel microtubules make the tracks for fast axonal transport. In dendrites, the microtubules are of mixed polarity, with some plus ends pointing outward and some pointing inward.

from the cell body all the way to the axon terminals; each is typically only a few micrometers in length, but large numbers are staggered in an overlapping array. This set of aligned microtubule tracks acts as a highway to transport many specific proteins, protein-containing vesicles, and mRNAs to the axon terminals, where synapses must be constructed and maintained. The longest axon in the human body reaches from the base of the spinal cord to the foot, being up to a meter in length.

Mitochondria, large numbers of specific proteins in transport vesicles, and synaptic vesicle precursors make the long journey in the forward (anterograde) direction. They are carried there by plus-end-directed kinesin-family motor proteins that can move them a meter in as little as two or three days, which is a great improvement over diffusion, which would take approximately several decades to move a mitochondrion this distance. Many members of the kinesin superfamily contribute to this *anterograde axonal transport*, most carrying specific subsets of membrane-enclosed organelles along the microtubules. The great diversity of the kinesin family motor proteins used in axonal transport suggests that they are involved in targeting their cargo to specific structures near the terminus or along the way, as well as in cargo movement. Old components from the axon terminals are carried back to the cell body for degradation and recycling by a *retrograde axonal transport*. This transport occurs along the same set of oriented microtubules, but it relies on cytoplasmic dynein, which is a minus-end-directed motor protein. Retrograde transport is also critical for communicating the presence of growth and survival signals received by the nerve terminus back to the nucleus, in order to influence gene expression.

One form of human peripheral neuropathy, Charcot-Marie-Tooth disease, is caused by a point mutation in a particular kinesin family member that transports synaptic vesicle precursors down the axon. Other kinds of neurodegenerative diseases such as Alzheimer's disease may also be caused in part by disruptions in neuronal trafficking; as pointed out previously, the amyloid precursor protein APP is part of a protein complex that serves as a receptor for kinesin-1 binding to other axonal transport vesicles.

Axonal structure depends on the axonal microtubules, as well as on the contributions of the other two major cytoskeletal systems—actin filaments and intermediate filaments. Actin filaments line the cortex of the axon, just beneath the plasma membrane, and actin-based motor proteins such as myosin V are also abundant in the axon, presumably to help move materials. Neurofilaments, the specialized intermediate filaments of nerve cells, provide the most important structural support in the axon. A disruption in neurofilament structure, or in the cross-linking proteins that attach the neurofilaments to the microtubules and actin filaments distributed along the axon, can result in axonal disorganization and eventually axonal degeneration.

The construction of the elaborate branching architecture of the neuron during embryonic development requires actin-based motility. As mentioned earlier, the tips of growing axons and dendrites extend by means of a *growth cone*, a specialized motile structure rich in actin (**Figure 16–105**). Most neuronal growth cones produce filopodia, and some make lamellipodia as well. The protrusion

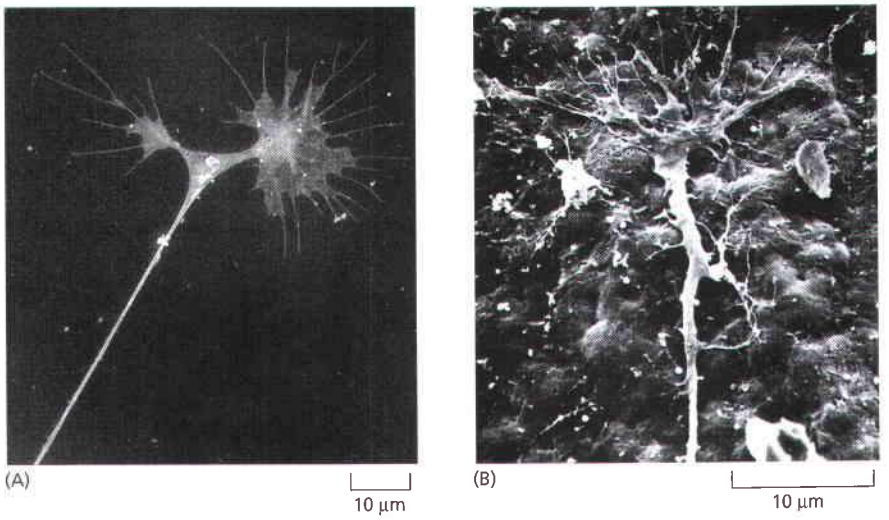


Figure 16–105 Neuronal growth cones. <AAGA> (A) Scanning electron micrograph of two growth cones at the end of a neurite, put out by a chick sympathetic neuron in culture. Here, a previously single growth cone has recently split into two. Note the many filopodia and the large lamellipodia. The taut appearance of the neurite is due to tension generated by the forward movement of the growth cones, which are often the only firm points of attachment of the axon to the substratum. (B) Scanning electron micrograph of the growth cone of a sensory neuron crawling over the inner surface of the epidermis of a *Xenopus* tadpole. (A, from D. Bray, in Cell Behaviour [R. Bellairs, A. Curtis and G. Dunn, eds.]. Cambridge, UK: Cambridge University Press, 1982; B, from A. Roberts, *Brain Res.* 118:526–530, 1976. With permission from Elsevier.)

and stabilization of growth-cone filopodia are exquisitely sensitive to environmental cues. Some cells secrete soluble proteins such as netrin to attract or repel growth cones. These modulate the structure and motility of the growth cone cytoskeleton by altering the balance between Rac activity and Rho activity at the leading edge (see Figure 15–62). In addition, there are fixed guidance markers along the way, attached to the extracellular matrix or to the surfaces of cells. When a filopodium encounters such a "guidepost" in its exploration, it quickly forms adhesive contacts. It is thought that a myosin-dependent collapse of the actin meshwork in the unstabilized part of the growth cone then causes the developing axon to turn toward the guidepost.

Thus, a complex combination of positive and negative signals, both soluble and insoluble, accurately guide the growth cone to its final destination. Microtubules then reinforce the directional decisions made by the actin-rich protrusive structures at the leading edge of the growth cone. Microtubules from the axonal parallel array just behind the growth cone are constantly growing into the growth cone and shrinking back by dynamic instability. Adhesive guidance signals are somehow relayed to the dynamic microtubule ends, so that microtubules growing in the correct direction are stabilized against disassembly. In this way, a microtubule-rich axon is left behind, marking the path that the growth cone has traveled.

Dendrites are generally much shorter projections than axons, and they receive synaptic inputs rather than being specialized for sending signals like axons. The microtubules in dendrites all lie parallel to one another but their polarities are mixed, with some pointing their plus ends toward the dendrite tip, while others point back toward the cell body. Nevertheless, dendrites also form as the result of growth-cone activity. Therefore, it is the growth cones at the tips of axons and dendrites that create the intricate, highly individual morphology of each mature neuronal cell (**Figure 16–106**).

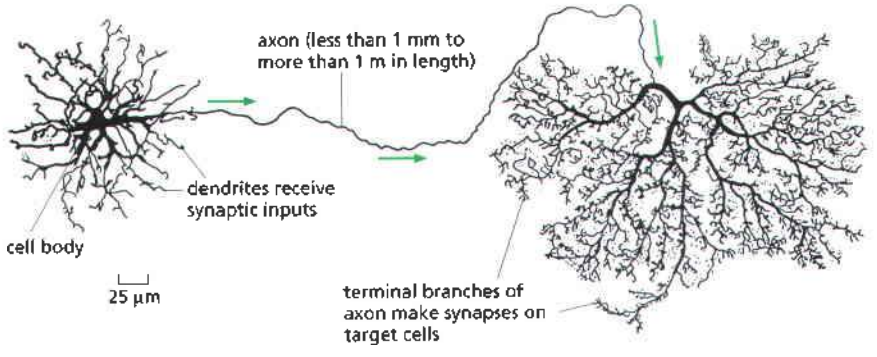


Figure 16–106 The complex architecture of a vertebrate neuron. The neuron shown is from the retina of a monkey. The arrows indicate the direction of travel of the electrical signal along the axon. The longest and largest neurons in the human body extend for a distance of about 1 m (1 million μm), from the base of the spinal cord to the tip of the big toe, and have an axon diameter of 15 μm. (Adapted from B.B. Boycott, in Essays on the Nervous System [R. Bellairs and E.G. Gray, eds.]. Oxford, UK: Clarendon Press, 1974.)

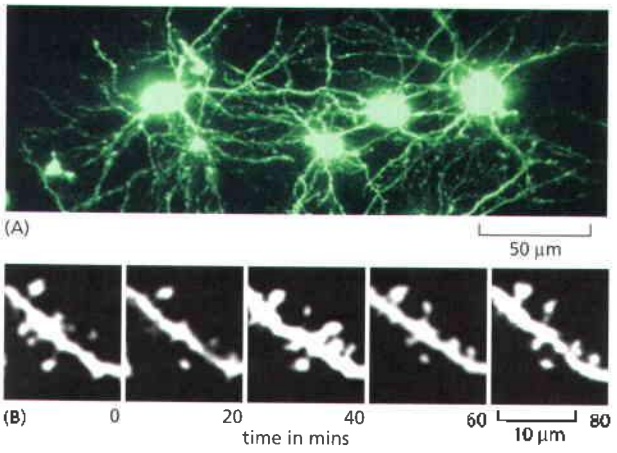


Figure 16–107 Rapid changes in dendrite structure within a living mouse brain. (A) Image of cortical neurons in a transgenic mouse that has been engineered to express green fluorescent protein in a small fraction of its brain cells. Changes in these brain neurons and their projections can be followed for months using highly sensitive fluorescence microscopy. To make this possible, the mouse is subjected to an operation that introduces a small transparent window through its skull, and it is anesthetized each time that an image is recorded. (B) A single dendrite, imaged over the period of 80 minutes, demonstrates that dendrites are constantly sending out and retracting tiny actin-dependent protrusions to create the dendritic spines that receive the vast majority of excitatory synapses from axons in the brain. Those spines that become stabilized and persist for months are thought to be important for brain function, and may be involved in long-term memory. (Courtesy of Karel Svoboda.)

Although the neurons of the central nervous system are long-lived cells, they are by no means static. Synapses are constantly being created, strengthened, weakened, and eliminated as the brain learns, evaluates, and forgets. High-resolution imaging of the structure of neurons in the brains of adult mice has revealed that neuronal morphology is undergoing constant rearrangement as synapses are forged and broken (**Figure 16–107**). These actin-dependent rearrangements are thought to be critical in learning and long-term memory. In this way, the cytoskeleton provides the engine for construction of the entire nervous system, as well as producing the supporting structures that strengthen, stabilize, and maintain its parts.

Summary

Two distinct types of specialized structures in eucaryotic cells are formed from highly ordered arrays of motor proteins that move on stabilized filament tracks. The myosin–actin system of the sarcomere powers the contraction of various types of muscle, including skeletal, smooth, and cardiac muscle. The dynein–microtubule system of the axoneme powers the beating of cilia and the undulations of flagella.

Whole-cell movements and the large-scale shaping and structuring of cells require the coordinated activities of all three basic filament systems along with a large variety of cytoskeletal accessory proteins, including motor proteins. During cell division, the functions of the microtubule-based mitotic spindle require spatial and temporal cooperation between dynamic cytoskeletal filaments, active molecular motor proteins, and a wide variety of accessory factors. Cell crawling—a widespread behavior important in embryonic development and also in wound healing, tissue maintenance, and immune system function in the adult animal—is another prime example of such complex, coordinated cytoskeletal action. For a cell to crawl, it must generate and maintain an overall structural polarity, which is influenced by external cues. In addition, the cell must coordinate protrusion at the leading edge (by assembly of new actin filaments), adhesion of the newly protruded part of the cell to the substratum, forces generated by molecular motors to bring the cell body forward.

Complex cells, such as neurons, require the coordinated assembly of microtubules, neurofilaments (neuronal intermediate filaments), and actin filaments, as well as the actions of dozens of highly specialized molecular motors that transport subcellular components to their appropriate destinations.

PROBLEMS

Which statements are true? Explain why or why not.

16–1 The role of ATP hydrolysis in actin polymerization is similar to the role of GTP hydrolysis in tubulin polymerization: both serve to weaken the bonds in the polymer and thereby promote depolymerization.

16–2 In most animal cells, minus end-directed microtubule motors deliver their cargo to the periphery of the cell, whereas plus end-directed microtubule motors deliver their cargo to the interior of the cell.

16–3 Motor neurons trigger action potentials in muscle cell membranes that open voltage-sensitive Ca^{2+} channels in T-tubules, allowing extracellular Ca^{2+} to enter the cytosol, bind to troponin C, and initiate rapid muscle contraction.

Discuss the following problems.

16–4 At 1.4 mg/ml, pure tubulin, microtubules grow at a rate of about 2 μm/min. At this growth rate how many αβ-tubulin dimers (8 nm in length) are added to the ends of a microtubule each second?

16–5 A solution of pure αβ-tubulin dimers is thought to nucleate microtubules by forming a linear protofilament about seven dimers in length. At that point, the probabilities that the next αβ-dimer will bind laterally or to the end of the protofilament are about equal. The critical event for microtubule formation is thought to be the first lateral association (Figure Q16–1). How does lateral association promote the subsequent rapid formation of a microtubule?

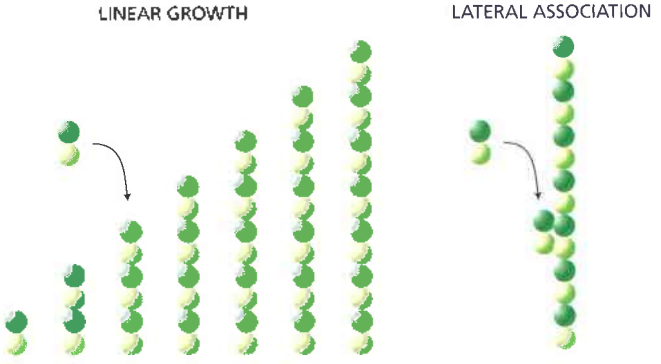


Figure Q16–1 Model for microtubule nucleation by pure αβ-tubulin dimers (Problem 16–5).

16–6 How does a centrosome "know" when it has found the center of the cell?

16–7 The concentration of actin in cells is 50–100 times greater than the critical concentration observed for pure actin in a test tube. How is this possible? What prevents the actin subunits in cells from polymerizing into filaments? Why is it advantageous to the cell to maintain such a large pool of actin subunits?

16–8 The movements of single motor-protein molecules can be analyzed directly. Using polarized laser light, it is possible to create interference patterns that exert a centrally directed force, ranging from zero at the center to a few piconewtons at the periphery (about 200 nm from the center). Individual molecules that enter the interference pattern are rapidly pushed to the center, allowing them to be captured and moved at the experimenter's discretion.

Using such "optical tweezers," single kinesin molecules can be positioned on a microtubule that is fixed to a coverslip. Although a single kinesin molecule cannot be seen optically, it can be tagged with a silica bead and tracked indirectly by following the bead (Figure Q16–2A). In the absence of ATP, the kinesin molecule remains at the center of the interference pattern, but with ATP it moves toward the plus end of the microtubule. As kinesin moves along the microtubule, it encounters the force of the interference pattern, which simulates the load kinesin carries during its actual function in the cell. Moreover, the pressure against the silica bead counters the effects of Brownian (thermal) motion, so that the position of the bead more accurately reflects the position of the kinesin molecule on the microtubule.

Traces of the movements of a kinesin molecule along a microtubule are shown in Figure Q16–2B.

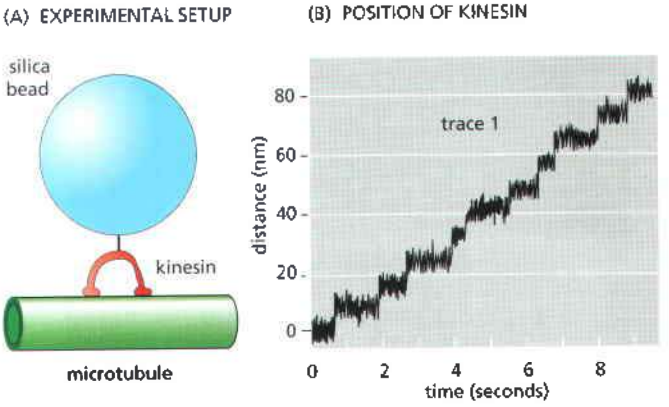


Figure Q16–2 Movement of kinesin along a microtubule (Problem 16–8). (A) Experimental setup with kinesin linked to a silica bead, moving along a microtubule. (B) Position of kinesin (as visualized by position of silica bead) relative to center of interference pattern, as a function of time of movement along the microtubule. The jagged nature of the trace results from Brownian motion of the bead.

A. As shown in Figure Q16–2B, all movement of kinesin is in one direction (toward the plus end of the microtubule). What supplies the free energy needed to ensure a unidirectional movement along the microtubule?

B. What is the average rate of movement of kinesin along the microtubule?

C. What is the length of each step that a kinesin takes as it moves along a microtubule?

D. From other studies it is known that kinesin has two globular domains that each can bind to β-tubulin, and that kinesin moves along a single protofilament in a microtubule. In each protofilament the β-tubulin subunit repeats at 8-nm intervals. Given the step length and the interval between β-tubulin subunits, how do you suppose a kinesin molecule moves along a microtubule?

E. Is there anything in the data in Figure Q16–2B that tells you how many ATP molecules are hydrolyzed per step?

16–9 How is the unidirectional motion of a lamellipodium maintained?

16–10 Detailed measurements of sarcomere length and tension during isometric contraction in striated muscle provided crucial early support for the sliding filament model of muscle contraction. Based on your understanding of the sliding filament model and the structure of a sarcomere, propose a molecular explanation for the relationship of tension to sarcomere length in the portions of **Figure Q16–3** marked I, II, III, and IV. (In this muscle, the length of the myosin filament is 1.6 μm and the lengths of the actin thin filaments that project from the Z discs are 1.0 μm.)

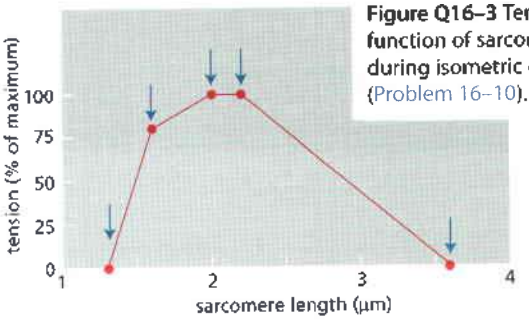


Figure Q16–3 Tension as a function of sarcomere length during isometric contraction (Problem 16–10).

REFERENCES

General

Bray D (2001) Cell Movements: From Molecules to Motility, 2nd ed. New York: Garland Science.

Howard J (2001) Mechanics of Motor Proteins and the Cytoskeleton. Sunderland, MA: Sinauer.

The Self-Assembly and Dynamic Structure of Cytoskeletal Filaments

Dogterom M & Yurke B (1997) Measurement of the force-velocity relation for growing microtubules. *Science* 278:856–860.

Garner EC, Campbell CS & Mullins RD (2004) Dynamic instability in a DNA-segregating prokaryotic actin homolog. *Science* 306:1021–1025.

Helfand BT, Chang L & Goldman RD (2003) The dynamic and motile properties of intermediate filaments. *Annu Rev Cell Dev Biol* 19:445–467.

Hill TL & Kirschner MW (1982) Bioenergetics and kinetics of microtubule and actin filament assembly-disassembly. *Int Rev Cytol* 78:1–125.

Hotani H & Horio T (1988) Dynamics of microtubules visualized by darkfield microscopy: treadmilling and dynamic instability. *Cell Motil Cytoskeleton* 10:229–236.

Jones LJ, Carballido-Lopez R & Errington J (2001) Control of cell shape in bacteria: helical, actin-like filaments in *Bacillus subtilis*. *Cell* 104:913–922.

Kerssemakers JW, Munteanu EL, Laan L et al (2006) Assembly dynamics of microtubules at molecular resolution. *Nature* 442:709–712.

Luby-Phelps K (2000) Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. *Int Rev Cytol* 192:189–221.

Mitchison T & Kirschner M (1984) Dynamic instability of microtubule growth. *Nature* 312:237–242.

Mitchison TJ (1995) Evolution of a dynamic cytoskeleton. *Philos Trans R Soc Lond B Biol Sci* 349:299–304.

Mukherjee A & Lutkenhaus J (1994) Guanine nucleotide-dependent assembly of FtsZ into filaments. *J Bacteriol* 176:2754–2758.

Oosawa F & Asakura S (1975) Thermodynamics of the polymerization of protein. New York: Academic Press, pp 41–55, pp 90–108.

Pauling L (1953) Aggregation of Globular Proteins. *Discuss Faraday Soc* 13:170–176.

Rodionov VI & Borisy GG (1997) Microtubule treadmilling *in vivo*. *Science* 275:215–218.

Shih YL & Rothfield L (2006) The bacterial cytoskeleton. *Microbiol Mol Biol Rev* 70:729–754.

Theriot JA (2000) The polymerization motor. *Traffic* 1:19–28.

How Cells Regulate Their Cytoskeletal Filaments

Aldaz H, Rice LM, Stearns T & Agard DA (2005) Insights into microtubule nucleation from the crystal structure of human gamma-tubulin. *Nature* 435:523–527.

Bretscher A, Chambers D, Nguyen R & Reczek D (2000) ERM-Merlin and EBP50 protein families in plasma membrane organization and function. *Annu Rev Cell Dev Biol* 16:113–143.

Doxsey S, McCollum D & Theurkauf W (2005) Centrosomes in cellular regulation. *Annu Rev Cell Dev Biol* 21:411–434.

Garcia ML & Cleveland DW (2001) Going new places using an old MAP: tau, microtubules and human neurodegenerative disease. *Curr Opin Cell Biol* 13:41–48.

Holy TE, Dogterom M, Yurke B & Leibler S (1997) Assembly and positioning of microtubule asters in microfabricated chambers. *Proc Natl Acad Sci USA* 94:6228–6231.

Mullins RD, Heuser JA & Pollard TD (1998) The interaction of Arp2/3 complex with actin: nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *Proc Natl Acad Sci USA* 95:6181–6186.

Robinson RC, Turbedsky K, Kaiser DA et al (2001) Crystal structure of Arp2/3 complex. *Science* 294:1679–1684.

Quarmby L (2000) Cellular Samurai: katanin and the severing of microtubules. *J Cell Sci* 113:2821–2827.

Stearns T & Kirschner M (1994) *In vitro* reconstitution of centrosome assembly and function: the central role of gamma-tubulin. *Cell* 76:623–637.

Wiese C & Zheng Y (2006) Microtubule nucleation: gamma-tubulin and beyond. *J Cell Sci* 119:4143–4153.

Zheng Y, Wong ML, Alberts B & Mitchison T (1995) Nucleation of microtubule assembly by a gamma-tubulin-containing ring complex. *Nature* 378:578–583.

Zigmond SH (2004) Formin-induced nucleation of actin filaments. *Curr Opin Cell Biol* 16:99–105.

Molecular Motors

Burgess SA, Walker ML, Sakakibara H et al (2003) Dynein structure and power stroke. *Nature* 421:715–718.

Hirokawa N (1998) Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science* 279:519–526.

Howard J, Hudspeth AJ & Vale RD (1989) Movement of microtubules by single kinesin molecules. *Nature* 342:154–158.

Howard J (1997) Molecular motors: structural adaptations to cellular functions. *Nature* 389:561–567.

Rayment I, Rypniewski WR, Schmidt-Base K et al (1993) Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science* 261:50–58.

Reck-Peterson SL, Yildiz A, Carter AP et al (2006) Single-molecule analysis of dynein processivity and stepping behavior. *Cell* 126:335–348.

Rice S, Lin AW, Safer D et al (1999) A structural change in the kinesin motor protein that drives motility. *Nature* 402:778–784.

Richards TA & Cavalier-Smith T (2005) Myosin domain evolution and the primary divergence of eukaryotes. *Nature* 436:1113–1118.

Svoboda K, Schmidt CF, Schnapp BJ & Block SM (1993) Direct observation of kinesin stepping by optical trapping interferometry. *Nature* 365:721–727.

Vikstrom KL & Leinwand LA (1996) Contractile protein mutations and heart disease. *Curr Opin Cell Biol* 8:97–105.

Wells AL, Lin AW, Chen LQ et al (1999) Myosin VI is an actin-based motor that moves backwards. *Nature* 401:505–508.

Yildiz A, Forkey JN, McKinney SA et al (2003) Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science* 300:2061–2065.

Yildiz A & Selvin PR (2005) Kinesin: walking, crawling or sliding along? *Trends Cell Biol* 15:112–120.

The Cytoskeleton and Cell Behavior

Abercrombie M (1980) The crawling movement of metazoan cells. *Proc Roy Soc B* 207:129–147.

Cooke R (2004) The sliding filament model: 1972–2004. *J Gen Physiol* 123:643–656.

Dent EW & Gertler FB (2003) Cytoskeletal dynamics and transport in growth cone motility and axon guidance. *Neuron* 40:209–227.

Lauffenburger DA & Horwitz AF (1996) Cell migration: a physically integrated molecular process. *Cell* 84:359–369.

Lo CM, Wang HB, Dembo M & Wang YL (2000) Cell movement is guided by the rigidity of the substrate. *Biophys J* 79:144–152.

Madden K & Snyder M (1998) Cell polarity and morphogenesis in budding yeast. *Annu Rev Microbiol* 52:687–744.

Ridley AJ, Schwartz MA, Burridge K et al (2003) Cell migration: integrating signals from front to back. *Science* 302:1704–1709.

Rafelski SM & Theriot JA (2004) Crawling toward a unified model of cell motility: spatial and temporal regulation of actin dynamics. *Annu Rev Biochem* 73:209–239.

Parent CA & Devreotes PN (1999) A cell's sense of direction. *Science* 284:765–770.

Pollard TD & Borisy GG (2003) Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112:453–465.

Purcell EM (1977) Life at low Reynolds' number. *Am J Phys* 45:3–11.

Wittmann T, Hyman A & Desai A (2001) The spindle: a dynamic assembly of microtubules and motors. *Nature Cell Biol* 3:E28–E34.

17

The Cell Cycle

In This Chapter

OVERVIEW OF THE CELL CYCLE 1054

THE CELL-CYCLE CONTROL SYSTEM 1060

S PHASE 1067

MITOSIS 1071

CYTOKINESIS 1092

CONTROL OF CELL DIVISION AND CELL GROWTH 1101

The only way to make a new cell is to duplicate a cell that already exists. This simple fact, first established in the middle of the nineteenth century, carries with it a profound message for the continuity of life. All living organisms, from the unicellular bacterium to the multicellular mammal, are products of repeated rounds of cell growth and division extending back in time to the beginnings of life on Earth over three billion years ago.

A cell reproduces by performing an orderly sequence of events in which it duplicates its contents and then divides in two. This cycle of duplication and division, known as the **cell cycle**, is the essential mechanism by which all living things reproduce. In unicellular species, such as bacteria and yeasts, each cell division produces a complete new organism. In multicellular species, long and complex sequences of cell divisions are required to produce a functioning organism. Even in the adult body, cell division is usually needed to replace cells that die. In fact, each of us must manufacture many millions of cells every second simply to survive: if all cell division were stopped—by exposure to a very large dose of x-rays, for example—we would die within a few days.

The details of the cell cycle vary from organism to organism and at different times in an organism's life. Certain characteristics, however, are universal. The minimum set of processes that a cell has to perform are those that allow it to accomplish its most fundamental task: the passing on of its genetic information to the next generation of cells. To produce two genetically identical daughter cells, the DNA in each chromosome must first be faithfully replicated to produce two complete copies, and the replicated chromosomes must then be accurately distributed *(segregated)* to the two daughter cells, so that each receives a copy of the entire genome (**Figure 17–1**).

Eucaryotic cells have evolved a complex network of regulatory proteins, known as the *cell-cycle control system*, that governs progression through the cell cycle. The core of this system is an ordered series of biochemical switches that initiate the main events of the cycle, including chromosome duplication and segregation. In most cells, additional layers of regulation enhance the fidelity of cell division and allow the control system to respond to various signals from both inside and outside the cell. Inside the cell, the control system monitors progression through the cell cycle and delays later events until earlier events have been completed. It does not permit preparations for the segregation of duplicated chromosomes, for example, until DNA replication is complete. The control system also monitors conditions outside the cell. In a multicellular animal, the system is highly responsive to signals from other cells, stimulating cell division when more cells are needed and blocking it when they are not. The cell-cycle control system therefore has a central role in regulating cell numbers in the tissues of the body. When the system malfunctions, excessive cell divisions can result in cancer.

In addition to duplicating their genome, most cells also duplicate their other organelles and macromolecules; otherwise, daughter cells would get smaller with each division. To maintain their size, dividing cells must coordinate their growth (that is, their increase in cell mass) with their division.

This chapter describes the various events of the cell cycle and how they are controlled and coordinated. We begin with a brief overview of the cell cycle. We then describe the cell-cycle control system and explain how it triggers the

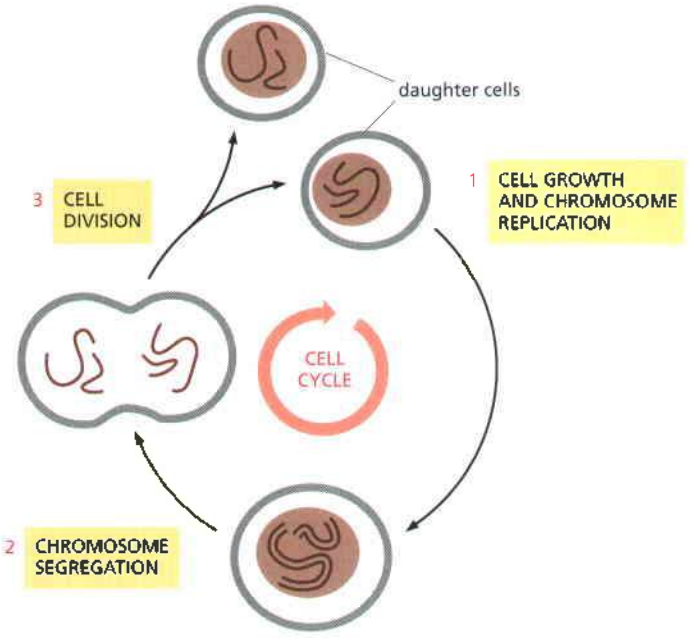


Figure 17–1 **The cell cycle.** The division of a hypothetical eucaryotic cell with two chromosomes is shown to illustrate how two genetically identical daughter cells are produced in each cycle. Each of the daughter cells will often continue to divide by going through additional cell cycles.

different events of the cycle. We next consider in detail the major stages of the cell cycle, in which the chromosomes are duplicated and then segregated into the two daughter cells. Finally, we consider how extracellular signals govern the rates of cell growth and division and how these two processes are coordinated.

OVERVIEW OF THE CELL CYCLE

We begin this section with a brief description of the four *phases* of the eucaryotic cell cycle. We then consider some of the methods and model cell systems used to study the cell cycle.

The Eucaryotic Cell Cycle Is Divided into Four Phases

The most basic function of the cell cycle is to duplicate accurately the vast amount of DNA in the chromosomes and then segregate the copies precisely into two genetically identical daughter cells. These processes define the two major phases of the cell cycle. Chromosome duplication occurs during *S phase* (S for DNA *s*ynthesis), which requires 10–12 hours and occupies about half of the cell-cycle time in a typical mammalian cell. After S phase, chromosome segregation and cell division occur in *M phase* (M for *m*itosis), which requires much less time (less than an hour in a mammalian cell). M phase comprises two major events: nuclear division, or *mitosis*, during which the copied chromosomes are distributed into a pair of daughter nuclei; and cytoplasmic division, or *cytokinesis*, when the cell itself divides in two (Figure 17–2).

At the end of S phase, the DNA molecules in each pair of duplicated chromosomes are intertwined and held tightly together by specialized protein linkages. Early in mitosis at a stage called *prophase*, the two DNA molecules are gradually disentangled and condensed into pairs of rigid and compact rods called **sister chromatids**, which remain linked together by *sister-chromatid cohesion*. When the nuclear envelope disassembles later in mitosis, the sister chromatid pairs become attached to the *mitotic spindle*, a giant bipolar array of microtubules (discussed in Chapter 16). Sister chromatids are attached to opposite poles of the spindle, and, eventually, all sisters align at the spindle equator in a stage called *metaphase*. The destruction of sister-chromatid cohesion at the start of *anaphase* separates the sister chromatids, which are pulled to opposite poles of the spindle. The spindle is then disassembled, and the segregated

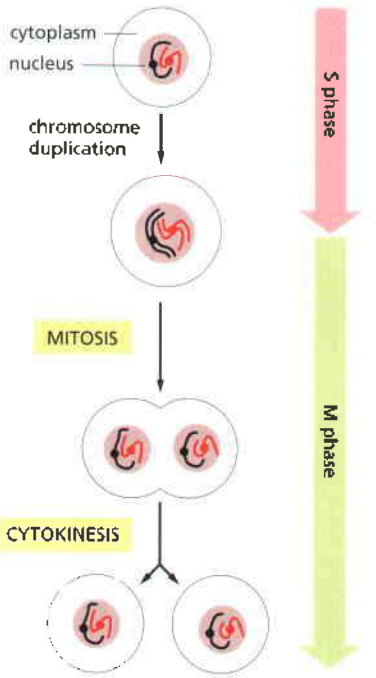
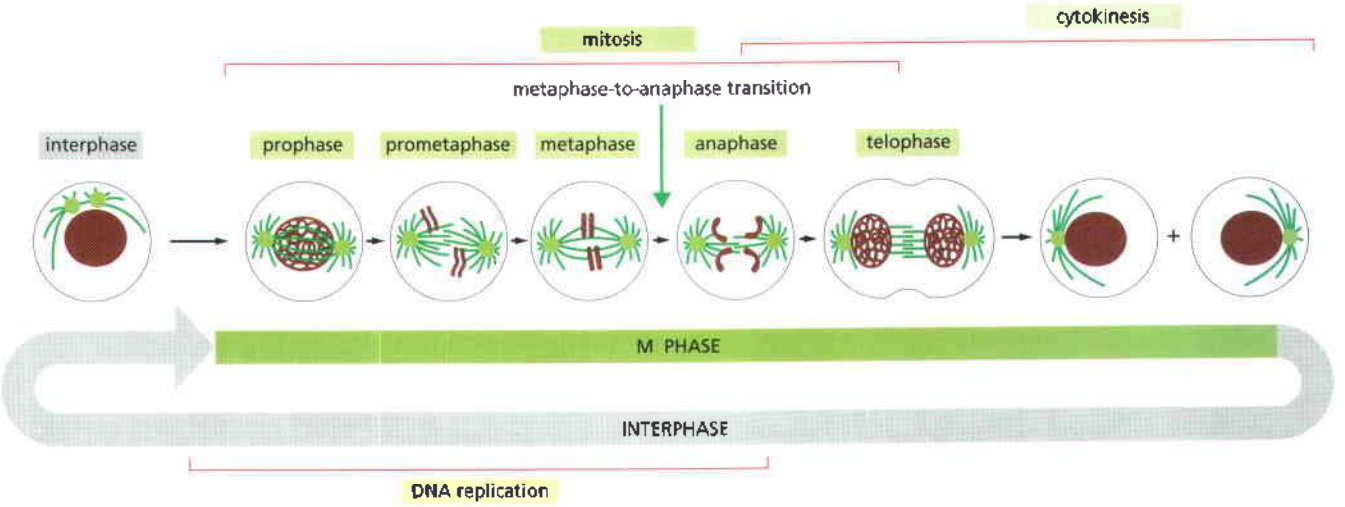


Figure 17–2 **The major events of the cell cycle.** The major chromosomal events of the cell cycle occur in S phase, when the chromosomes are duplicated, and M phase, when the duplicated chromosomes are segregated into a pair of daughter nuclei (in mitosis), after which the cell itself divides into two (cytokinesis).



chromosomes are packaged into separate nuclei at *telophase*. Cytokinesis then cleaves the cell in two, so that each daughter cell inherits one of the two nuclei (Figure 17–3). <TACT> <TCAA>

Most cells require much more time to grow and double their mass of proteins and organelles than they require to duplicate their chromosomes and divide. Partly to allow more time for growth, most cell cycles have extra *gap phases*—a **G_1 phase** between M phase and S phase and a **G_2 phase** between S phase and mitosis. Thus, the eucaryotic cell cycle is traditionally divided into four sequential phases: G_1, S, G_2, and M. G_1, S, and G_2 together are called **interphase** (Figure 17–4, and see Figure 17–3). In a typical human cell proliferating in culture, interphase might occupy 23 hours of a 24-hour cycle, with 1 hour for M phase. Cell growth occurs throughout the cell cycle, except during mitosis.

The two gap phases are more than simple time delays to allow cell growth. They also provide time for the cell to monitor the internal and external environment to ensure that conditions are suitable and preparations are complete before the cell commits itself to the major upheavals of S phase and mitosis. The G_1 phase is especially important in this respect. Its length can vary greatly depending on external conditions and extracellular signals from other cells. If extracellular conditions are unfavorable, for example, cells delay progress through G_1 and may even enter a specialized resting state known as G_0 (G zero), in which they can remain for days, weeks, or even years before resuming proliferation. Indeed, many cells remain permanently in G_0 until they or the organism dies. If extracellular conditions are favorable and signals to grow and divide are present, cells in early G_1 or G_0 progress through a commitment point near the end of G_1 known as **Start** (in yeasts) or the **restriction point** (in mammalian cells). We will use the term Start for both yeast and animal cells. After passing this point, cells are committed to DNA replication, even if the extracellular signals that stimulate cell growth and division are removed.

Figure 17–3 **The events of eucaryotic cell division as seen under a microscope.** The easily visible processes of nuclear division (mitosis) and cell division (cytokinesis), collectively called M phase, typically occupy only a small fraction of the cell cycle. The other, much longer, part of the cycle is known as interphase, which includes S phase and the gap phases (discussed in text). The five stages of mitosis are shown: an abrupt change in the biochemical state of the cell occurs at the transition from metaphase to anaphase. A cell can pause in metaphase before this transition point, but once it passes this point, the cell carries on to the end of mitosis and through cytokinesis into interphase.

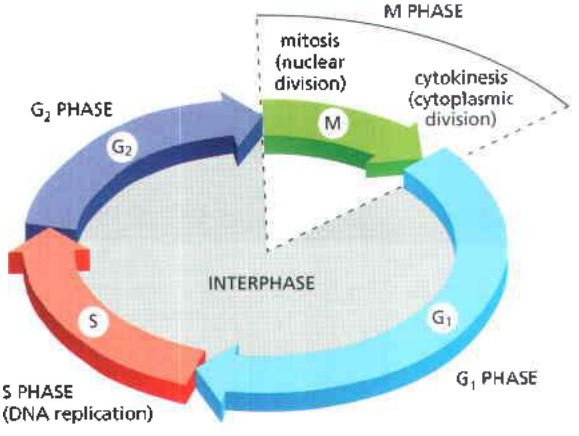


Figure 17–4 **The four phases of the cell cycle.** In most cells, gap phases separate the major events of S phase and M phase. G_1 is the gap between M phase and S phase, while G_2 is the gap between S phase and M phase.

Cell-Cycle Control Is Similar in All Eucaryotes

Some features of the cell cycle, including the time required to complete certain events, vary greatly from one cell type to another, even in the same organism. The basic organization of the cycle, however, is essentially the same in all eucaryotic cells, and all eucaryotes appear to use similar machinery and control mechanisms to drive and regulate cell-cycle events. The proteins of the cell-cycle control system, for example, first appeared over a billion years ago. Remarkably, they have been so well conserved over the course of evolution that many of them function perfectly when transferred from a human cell to a yeast cell. We can therefore study the cell cycle and its regulation in a variety of organisms and use the findings from all of them to assemble a unified picture of how eucaryotic cells divide. In the rest of this section, we briefly review the three eucaryotic systems most commonly used to study cell-cycle organization and control: yeasts, animal embryos, and cultured mammalian cells.

Cell-Cycle Control Can Be Dissected Genetically by Analysis of Yeast Mutants

Yeasts are tiny, single-celled fungi, with a cell-cycle control system remarkably similar to our own. Two species are generally used in studies of the cell cycle. The **fission yeast** *Schizosaccharomyces pombe* is named after the African beer it is used to produce. It is a rod-shaped cell that grows by elongation at its ends. Division occurs when a septum, or cell plate, forms midway along the rod (**Figure 17–5A**). The **budding yeast** *Saccharomyces cerevisiae* is used by both brewers and bakers. It is an oval cell that divides by forming a bud, which first appears during G_1 and grows steadily until it separates from the mother cell after mitosis (Figure 17–5B).

Despite their outward differences, the two yeast species share many features that are extremely useful for genetic studies. They reproduce almost as rapidly as bacteria and have a genome size less than 1% that of a mammal. They are amenable to rapid molecular genetic manipulation, in which genes can be deleted, replaced, or altered. Most importantly, they have the ability to proliferate in a *haploid* state, with only a single copy of each gene present in the cell. When cells are haploid, it is easy to isolate and study mutations that inactivate a gene, because we avoid the complication of having a second copy of the gene in the cell.

Many important discoveries about cell-cycle control have come from systematic searches for mutations in yeasts that inactivate genes encoding essential components of the cell-cycle control system. The genes affected by some of

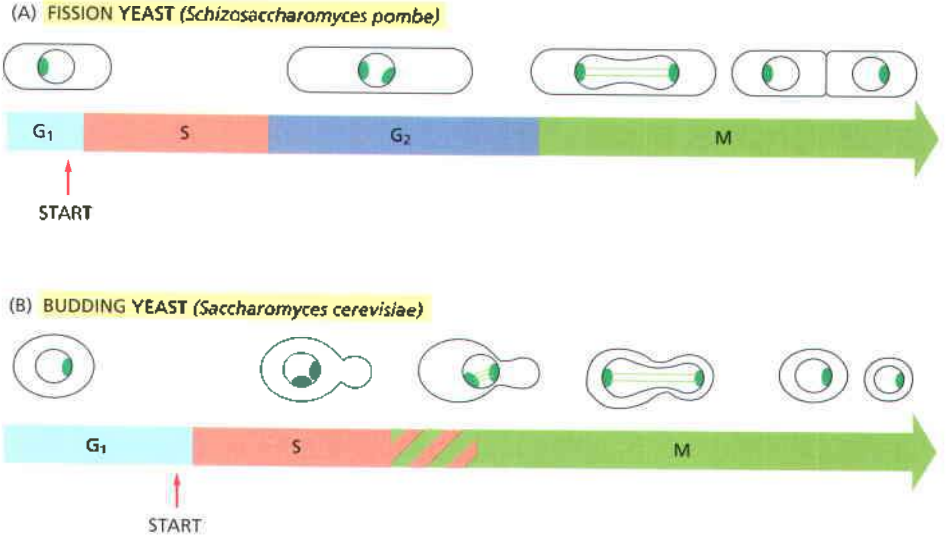


Figure 17–5 A comparison of the cell cycles of fission yeasts and budding yeasts. (A) The fission yeast has a typical eucaryotic cell cycle with G_1, S, G_2, and M phases. The nuclear envelope of the yeast cell, unlike that of a higher eucaryotic cell, does not break down during M phase. The microtubules of the mitotic spindle *(light green)* form inside the nucleus and are attached to spindle pole bodies *(dark green)* at its periphery. The cell divides by forming a partition (known as the cell plate) and splitting in two. (B) The budding yeast has normal G_1 and S phases but does not have a normal G_2 phase. Instead, a microtubule-based spindle begins to form late in S phase; as in fission yeasts, the nuclear envelope remains intact during mitosis, and the spindle forms within the nucleus. In contrast with a fission yeast cell, the cell divides by budding.

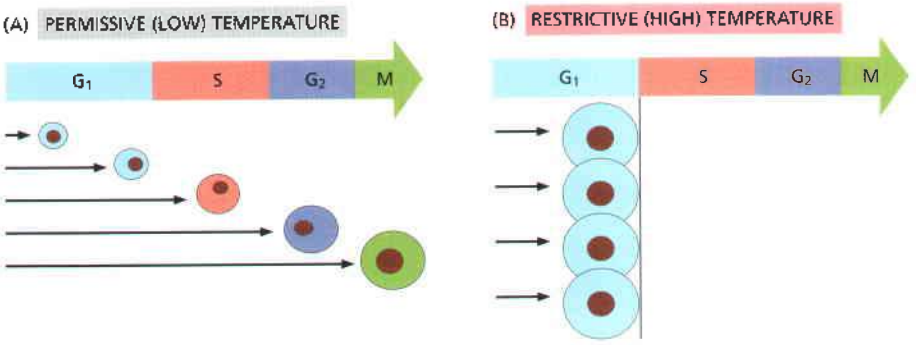


Figure 17–6 **The behavior of a temperature-sensitive *Cdc* mutant.** (A) At the permissive (low) temperature, the cells divide normally and are found in all phases of the cycle (the phase of the cell is indicated by its color). (B) On warming to the restrictive (high) temperature, at which the mutant gene product functions abnormally, the mutant cells continue to progress through the cycle until they come to the specific step that they are unable to complete (initiation of S phase, in this example). Because the *Cdc* mutants still continue to grow, they become abnormally large. By contrast, non-*Cdc* mutants, if deficient in a process that is necessary throughout the cycle for biosynthesis and growth (such as ATP production), halt haphazardly at any stage of the cycle—depending on when their biochemical reserves run out (not shown).

these mutations are known as **cell-division-cycle genes**, or ***Cdc* genes**. Many of these mutations cause cells to arrest at a specific point in the cell cycle, suggesting that the normal gene product is required to get the cell past this point.

A mutant that cannot complete the cell cycle, however, cannot be propagated. Thus, *Cdc* mutants can be selected and maintained only if their phenotype is *conditional*—that is, if the mutant gene product fails to function only in certain specific conditions. Most conditional cell-cycle mutations are *temperature-sensitive mutations*, in which the mutant protein fails to function at high temperatures but functions well enough to allow cell division at low temperatures. A temperature-sensitive *Cdc* mutant can be propagated at a low temperature (the *permissive condition*) and then raised to a higher temperature (the *restrictive condition*) to switch off the function of the mutant gene. At the higher temperature, cells continue through the cell cycle until they reach the point where the function of the mutant gene is required for further progress, and at this point they halt (Figure 17–6). In budding yeasts, we can detect a uniform cell-cycle arrest of this type by just looking at the cells: the presence or absence of a bud, and bud size, indicate the point in the cycle at which the mutant is arrested (Figure 17–7).

Cell-Cycle Control Can Be Analyzed Biochemically in Animal Embryos

The biochemical features of the cell cycle are readily analyzed in the giant fertilized eggs of many animals, which carry large stockpiles of the proteins needed for cell division. The egg of the frog *Xenopus*, for example, is over 1 mm in diameter and contains 100,000 times more cytoplasm than an average cell in the human body (Figure 17–8). Fertilization of the *Xenopus* egg triggers an astonishingly rapid sequence of cell divisions, called *cleavage divisions*, in which the single giant cell divides, without growing, to generate an embryo containing

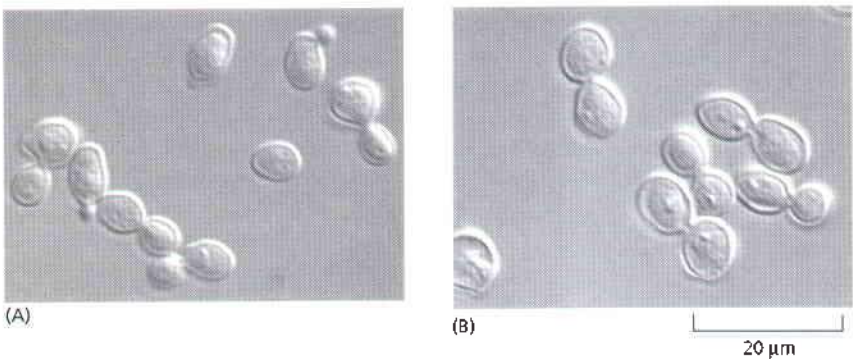


Figure 17–7 **The morphology of budding yeast cells arrested by a *Cdc* mutation.** (A) In a normal population of proliferating yeast cells, buds vary in size according to the cell-cycle stage. (B) In a *Cdc15* mutant grown at the restrictive temperature, cells complete anaphase but cannot complete the exit from mitosis and cytokinesis. As a result, they arrest uniformly with large buds, which are characteristic of late M phase. (Courtesy of Jeff Ubersax.)

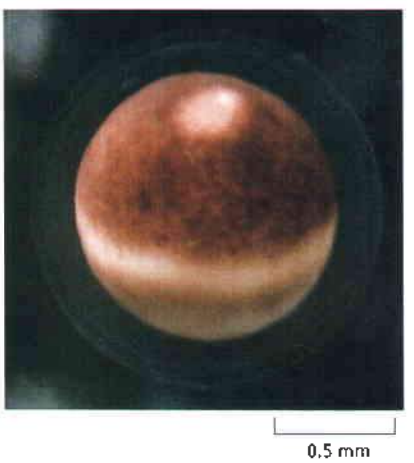


Figure 17–8 **A mature *Xenopus* egg, ready for fertilization.** The pale spot near the top shows the site of the nucleus, which has displaced the *brown* pigment in the surface layer of the egg cytoplasm. Although this cannot be seen in the picture, the nuclear envelope has broken down during the process of egg maturation. (Courtesy of Tony Mills.)

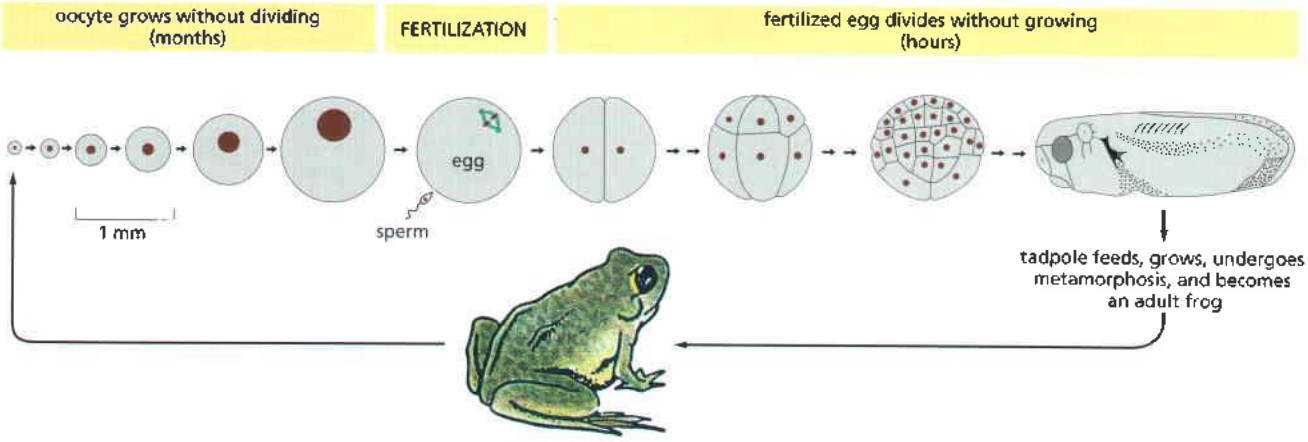


Figure 17–9 Oocyte growth and egg cleavage in *Xenopus*. The oocyte grows without dividing for many months in the ovary of the mother frog and finally matures into an egg (discussed in Chapter 21). Upon fertilization, the egg cleaves very rapidly—initially at a rate of one division cycle every 30 minutes—forming a multicellular tadpole within a day or two. The cells get progressively smaller with each division, and the embryo remains the same size. Growth starts only when the tadpole begins feeding. The drawings in the top row are all on the same scale (but the frog below is not).

thousands of smaller cells (**Figure 17–9**). <ATTT> After a first division that takes about 90 minutes, the next 11 divisions occur, more or less synchronously, at 30-minute intervals, producing about 4096 (2^{12}) cells within 7 hours. Each cycle is divided into S and M phases of about 15 minutes each, without detectable G_1 or G_2 phases.

The early embryonic cells of *Xenopus*, as well as those of the clam *Spisula* and the fruit fly *Drosophila*, are thus capable of exceedingly rapid division in the absence of either growth or many of the control mechanisms (discussed later) that operate in more complex cell cycles. These *early embryonic cell cycles* therefore reveal the workings of the cell-cycle control system stripped down and simplified to the minimum needed to achieve the most fundamental requirements—the duplication of the genome and its segregation into two daughter cells. Another advantage of these early embryos for cell-cycle analysis is their large size. It is relatively easy to inject test substances into an egg to determine their effect on cell-cycle progression. It is also possible to prepare almost pure cytoplasm from *Xenopus* eggs and reconstitute many events of the cell cycle in a test tube (**Figure 17–10**). In such cell extracts, we can observe and manipulate cell-cycle events under highly simplified and controllable conditions.

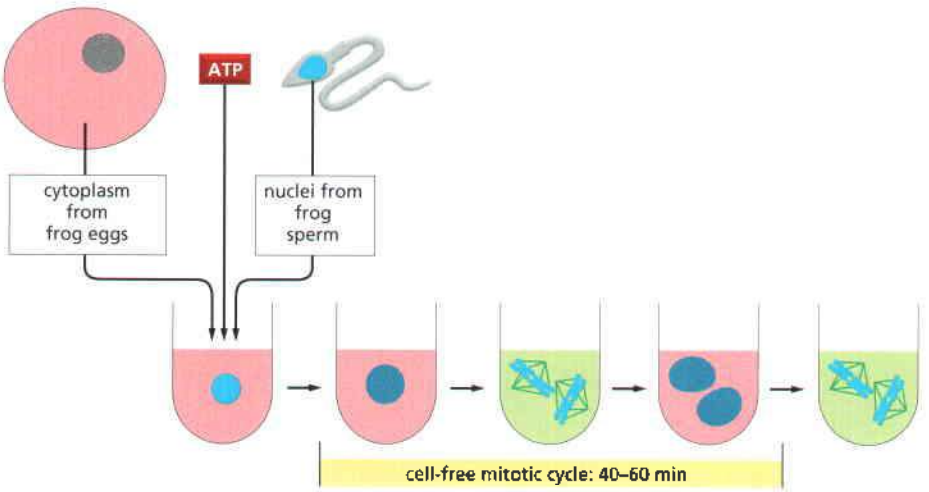


Figure 17–10 Studying the cell cycle in a cell-free system. Gentle centrifugation is used to break open a large batch of frog eggs and separate the cytoplasm from other cell components. The undiluted cytoplasm is collected, and sperm nuclei are added to it, together with ATP. The sperm nuclei decondense and then go through repeated cycles of DNA replication and mitosis, indicating that the cell-cycle control system is operating in this cell-free cytoplasmic extract.

Cell-Cycle Control Can Be Studied in Cultured Mammalian Cells

It is not easy to observe individual cells in an intact mammal. Most studies on mammalian cell-cycle control therefore use cells that have been isolated from normal tissues or tumors and grown in plastic culture dishes in the presence of essential nutrients and other factors (**Figure 17–11**). There is a complication, however. When cells from normal mammalian tissues are cultured in standard conditions, they often stop dividing after a limited number of division cycles. Human fibroblasts, for example, permanently cease dividing after 25–40 divisions, a process called *replicative cell senescence*, as we discuss later.

Mammalian cells occasionally undergo mutations that help them proliferate indefinitely in culture as "immortalized" *cell lines*. Although they are not normal, such cell lines are widely used for cell-cycle studies—and for cell biology generally—because they provide an unlimited source of genetically homogeneous cells. In addition, these cells are sufficiently large to allow detailed cytological observations of cell-cycle events, and they are amenable to biochemical analysis of the proteins involved in cell-cycle control.

Studies of cultured mammalian cells have been especially useful for examining the molecular mechanisms governing the control of cell proliferation in multicellular organisms. Such studies are important not only for understanding the normal controls of cell numbers in tissues but also for understanding the loss of these controls in cancer (discussed in Chapter 20).

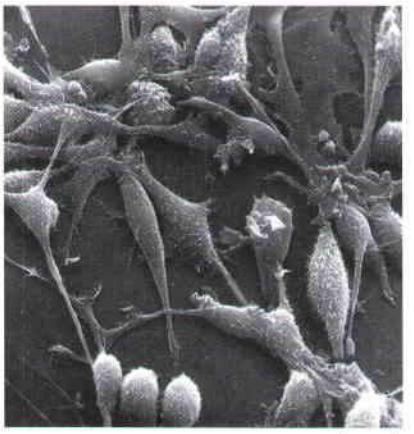


Figure 17–11 Mammalian cells proliferating in culture. The cells in this scanning electron micrograph are rat fibroblasts. (Courtesy of Guenter Albrecht-Buehler.)

Cell-Cycle Progression Can Be Studied in Various Ways

How can we tell what stage an animal cell has reached in the cell cycle? One way is simply to look at living cells with a microscope. A glance at a population of mammalian cells proliferating in culture reveals that a fraction of the cells have rounded up and are in mitosis (see Figure 17–11). Others can be observed in the process of cytokinesis. We can gain additional clues about cell-cycle position by staining cells with DNA-binding fluorescent dyes (which reveal the condensation of chromosomes in mitosis) or with antibodies that recognize specific cellular components such as the microtubules (revealing the mitotic spindle). Similarly, S-phase cells can be identified in the microscope by supplying them with visualizable molecules that are incorporated into newly synthesized DNA, such as the artificial thymidine analog bromo-deoxyuridine (BrdU). Cell nuclei that have incorporated BrdU are then visualized by staining with anti-BrdU antibodies (**Figure 17–12**).

Typically, in a population of cells that are all proliferating rapidly but asynchronously, about 30–40% will be in S phase at any instant and become labeled by a brief pulse of BrdU. From the proportion of cells in such a population that are labeled (the *labeling index*), we can estimate the duration of S phase as a fraction of the whole cell-cycle duration. Similarly, from the proportion of cells in mitosis (the *mitotic index*), we can estimate the duration of M phase. In addition, by giving a pulse of BrdU and allowing the cells to continue around the cycle for measured lengths of time, we can determine how long it takes for an S-phase cell to progress through G_2 into M phase, through M phase into G_1, and finally through G_1 back into S phase.

Another way to assess the stage that a cell has reached in the cell cycle is by measuring its DNA content, which doubles during S phase. This approach is greatly facilitated by the use of fluorescent DNA-binding dyes and a *flow cytometer*, which allows the rapid and automatic analysis of large numbers of cells (**Figure 17–13**). We can also use flow cytometry to determine the lengths of G_1, S,

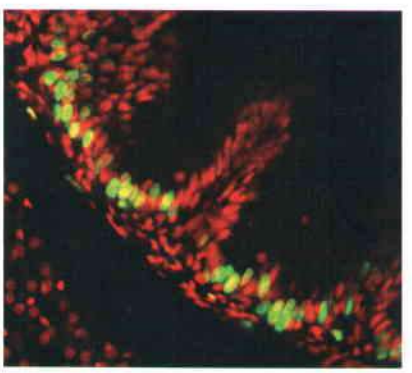


Figure 17–12 Labeling S-phase cells. An immunofluorescence micrograph of BrdU-labeled epithelial cells of the zebrafish gut. The fish was exposed to BrdU, after which the tissue was fixed and prepared for labeling with fluorescent anti-BrdU antibodies *(green)*. All the cells are stained with a *red* fluorescent dye. (Courtesy of Cécile Crosnier.)

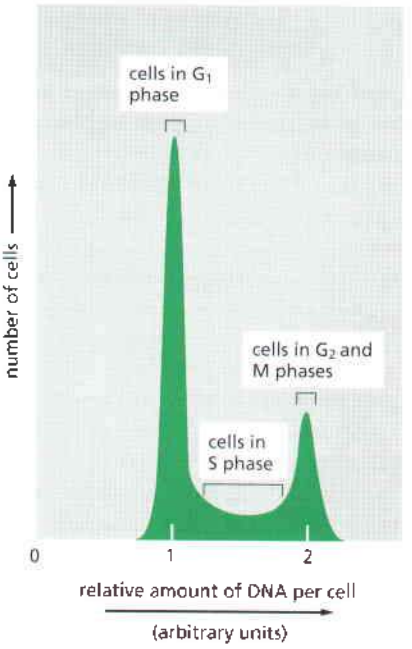


Figure 17–13 Analysis of DNA content with a flow cytometer. This graph shows typical results obtained for a proliferating cell population when the DNA content of its individual cells is determined in a flow cytometer. (A flow cytometer, also called a fluorescence-activated cell sorter, or FACS, can also be used to sort cells according to their fluorescence—see Figure 8–2). The cells analyzed here were stained with a dye that becomes fluorescent when it binds to DNA, so that the amount of fluorescence is directly proportional to the amount of DNA in each cell. The cells fall into three categories: those that have an unreplicated complement of DNA and are therefore in G_1, those that have a fully replicated complement of DNA (twice the G_1 DNA content) and are in G_2 or M phase, and those that have an intermediate amount of DNA and are in S phase. The distribution of cells in the case illustrated indicates that there are greater numbers of cells in G_1 than in G_2 + M phase, showing that G_1 is longer than G_2 + M in this population.

and G_2 + M phases, by following over time a population of DNA-labeled cells that have been preselected to be in one particular phase of the cell cycle: DNA content measurements on such a synchronized population of cells reveal how the cells progress through the cycle.

Summary

Cell division usually begins with duplication of the cell's contents, followed by distribution of those contents into two daughter cells. Chromosome duplication occurs during S phase of the cell cycle, whereas most other cell components are duplicated continuously throughout the cycle. During M phase, the replicated chromosomes are segregated into individual nuclei (mitosis), and the cell then splits in two (cytokinesis). S phase and M phase are usually separated by gap phases called G_1 and G_2, when various intracellular and extracellular signals regulate cell-cycle progression. Cell-cycle organization and control have been highly conserved during evolution, and studies in a wide range of systems—including yeasts, animal embryos, and mammalian cells in culture—have led to a unified view of eucaryotic cell-cycle control.

THE CELL-CYCLE CONTROL SYSTEM

For many years cell biologists watched the puppet show of DNA synthesis, mitosis, and cytokinesis but had no idea of what lay behind the curtain controlling these events. The cell-cycle control system was simply a black box inside the cell. It was not even clear whether there was a separate control system, or whether the processes of DNA synthesis, mitosis, and cytokinesis somehow controlled themselves. A major breakthrough came in the late 1980s with the identification of the key proteins of the control system, along with the realization that they are distinct from the proteins that perform the processes of DNA replication, chromosome segregation, and so on.

In this section, we first consider the basic principles upon which the cell-cycle control system operates. We then discuss the protein components of the control system and how they work together to time and coordinate the events of the cell cycle.

The Cell-Cycle Control System Triggers the Major Events of the Cell Cycle

The **cell-cycle control system** operates much like a timer or oscillator that triggers the events of the cell cycle in a set sequence (**Figure 17–14**). In its simplest form—as seen in the stripped-down embryonic cell cycles discussed earlier—the control system is like a rigidly programmed timer that provides a fixed amount of time for the completion of each cell-cycle event. The control system

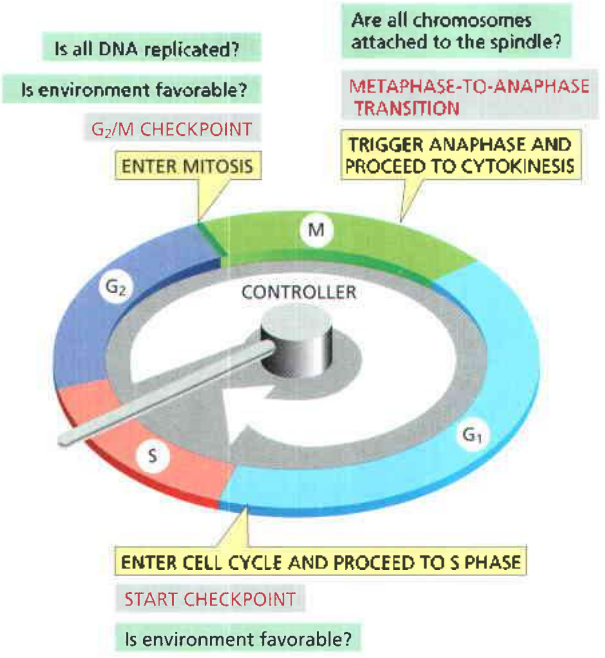


Figure 17–14 The control of the cell cycle. A cell-cycle control system triggers the essential processes of the cell cycle—such as DNA replication, mitosis, and cytokinesis. The control system is represented here as a central arm—the controller—that rotates clockwise, triggering essential processes when it reaches specific checkpoints on the outer dial. Information about the completion of cell-cycle events, as well as signals from the environment, can cause the control system to arrest the cycle at these checkpoints. The most prominent checkpoints occur at locations marked with *yellow boxes*.

in these cells is independent of the events it controls, so that its timing mechanisms continue to operate even if those events fail. In most cells, however, the control system does respond to information received back from the processes it controls. Sensors, for example, detect the completion of DNA synthesis, and if some malfunction prevents the successful completion of this process, signals are sent to the control system to delay progression to M phase. Such delays provide time for the machinery to be repaired and also prevent the disaster that might result if the cycle progressed prematurely to the next stage—and segregated incompletely replicated chromosomes, for example.

The cell-cycle control system is based on a connected series of biochemical switches, each of which initiates a specific cell-cycle event. This system of switches possesses many important engineering features that increase the accuracy and reliability of cell-cycle progression. First, the switches are generally *binary* (on/off) and launch events in a complete, irreversible fashion. It would clearly be disastrous, for example, if events like chromosome condensation or nuclear envelope breakdown were only partially initiated or started but not completed. Second, the cell-cycle control system is remarkably robust and reliable, partly because backup mechanisms and other features allow the system to operate effectively under a variety of conditions and even if some components fail. Finally, the control system is highly adaptable and can be modified to suit specific cell types or to respond to specific intracellular or extracellular signals.

In most eucaryotic cells, the cell-cycle control system triggers cell-cycle progression at three major regulatory transitions, or **checkpoints** (see Figure 17–14). The first checkpoint is Start (or the restriction point) in late G_1, where the cell commits to cell-cycle entry and chromosome duplication, as mentioned earlier. The second is the **G_2/M checkpoint**, where the control system triggers the early mitotic events that lead to chromosome alignment on the spindle in metaphase. The third is the **metaphase-to-anaphase transition**, where the control system stimulates sister-chromatid separation, leading to the completion of mitosis and cytokinesis. The control system blocks progression through each of these checkpoints if it detects problems inside or outside the cell. If the control system senses problems in the completion of DNA replication, for example, it will hold the cell at the G_2/M checkpoint until those problems are solved. Similarly, if extracellular conditions are not appropriate for cell proliferation, the control system blocks progression through Start, thereby preventing cell division until conditions become favorable.

The Cell-Cycle Control System Depends on Cyclically Activated Cyclin-Dependent Protein Kinases (Cdks)

Central components of the cell-cycle control system are members of a family of protein kinases known as **cyclin-dependent kinases (Cdks)**. The activities of these kinases rise and fall as the cell progresses through the cycle, leading to cyclical changes in the phosphorylation of intracellular proteins that initiate or regulate the major events of the cell cycle. An increase in Cdk activity at the G_2/M checkpoint, for example, increases the phosphorylation of proteins that control chromosome condensation, nuclear envelope breakdown, spindle assembly, and other events that occur at the onset of mitosis.

Cyclical changes in Cdk activity are controlled by a complex array of enzymes and other proteins that regulate these kinases. The most important of these Cdk regulators are proteins known as **cyclins**. Cdks, as their name implies, are dependent on cyclins for their activity: unless they are tightly bound to a cyclin, they have no protein kinase activity (Figure 17–15). Cyclins were originally named because they undergo a cycle of synthesis and degradation in each cell cycle. The levels of the Cdk proteins, by contrast, are constant, at least in the simplest cell cycles. Cyclical changes in cyclin protein levels result in the cyclic assembly and activation of the **cyclin–Cdk complexes**; this activation in turn triggers cell-cycle events.

There are four classes of cyclins, each defined by the stage of the cell cycle at which they bind Cdks and function. All eucaryotic cells require three of these classes (Figure 17–16):

1. **G_1/S-cyclins** activate Cdks in late G_1 and thereby help trigger progression through Start, resulting in a commitment to cell-cycle entry. Their levels fall in S phase.
2. **S-cyclins** bind Cdks soon after progression through Start and help stimulate chromosome duplication. S-cyclin levels remain elevated until mitosis, and these cyclins also contribute to the control of some early mitotic events.
3. **M-cyclins** activate Cdks that stimulate entry into mitosis at the G_2/M checkpoint. Mechanisms that we discuss later destroy M-cyclins in mid-mitosis.

In most cells, a fourth class of cyclins, the **G_1-cyclins**, helps govern the activities of the G_1/S cyclins, which control progression through Start in late G_1.

In yeast cells, a single Cdk protein binds all classes of cyclins and triggers different cell-cycle events by changing cyclin partners at different stages of the cycle. In vertebrate cells, by contrast, there are four Cdks. Two interact with G_1-cyclins, one with G_1/S- and S-cyclins, and one with M-cyclins. In this chapter, we simply refer to the different cyclin–Cdk complexes as **G_1-Cdk**, **G_1/S-Cdk**, **S-Cdk**, and **M-Cdk**. Table 17–1 lists the names of the individual Cdks and cyclins.

How do different cyclin–Cdk complexes trigger different cell-cycle events? The answer, at least in part, seems to be that the cyclin protein does not simply activate its Cdk partner but also directs it to specific target proteins. As a result,



Figure 17–15 Two key components of the cell-cycle control system. When cyclin forms a complex with Cdk, the protein kinase is activated to trigger specific cell-cycle events. Without cyclin, Cdk is inactive.

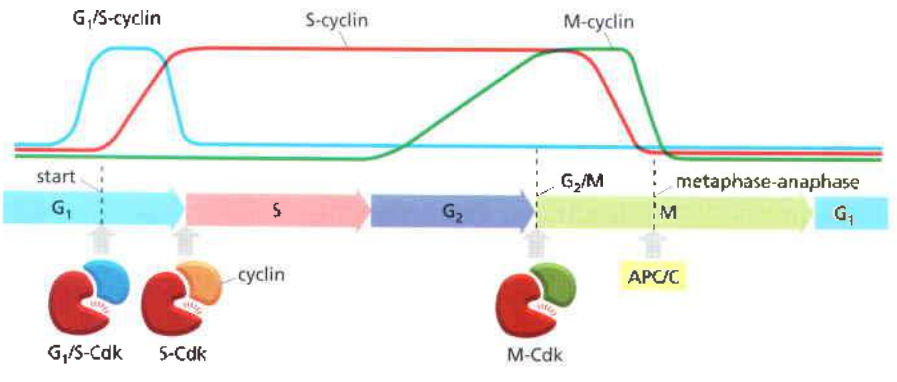


Figure 17–16 Cyclin–Cdk complexes of the cell-cycle control system. The concentrations of the three major cyclin types oscillate during the cell cycle, while the concentrations of Cdks (not shown) do not change and exceed the amounts of cyclins. In late G_1, rising G_1/S-cyclin levels lead to the formation of G_1/S-Cdk complexes that trigger progression through the Start checkpoint. S-Cdk complexes form at the start of S phase and trigger DNA replication, as well as some early mitotic events. M-Cdk complexes form during G_2 but are held in an inactive state by mechanisms we describe later. These complexes are activated at the end of G_2 and trigger the early events of mitosis. A separate regulatory protein, the APC/C, which we discuss later, initiates the metaphase-to-anaphase transition.

Table 17–1 The Major Cyclins and Cdks of Vertebrates and Budding Yeast

CYCLIN–CDK COMPLEX	VERTEBRATES		BUDDING YEAST	
	CYCLIN	CDK PARTNER	CYCLIN	CDK PARTNER
G_1-Cdk	cyclin D*	Cdk4, Cdk6	Cln3	Cdk1**
G_1/S-Cdk	cyclin E	Cdk2	Cln1, 2	Cdk1
S-Cdk	cyclin A	Cdk2, Cdk1**	Clb5, 6	Cdk1
M-Cdk	cyclin B	Cdk1	Clb1, 2, 3, 4	Cdk1

* There are three D cyclins in mammals (cyclins D1, D2, and D3).
** The original name of Cdk1 was Cdc2 in both vertebrates and fission yeast, and Cdc28 in budding yeast.

each cyclin–Cdk complex phosphorylates a different set of substrate proteins. The same cyclin–Cdk complex can also induce different effects at different times in the cycle, probably because the accessibility of some Cdk substrates changes during the cell cycle. Certain proteins that function in mitosis, for example, may become available for phosphorylation only in G_2.

Studies of the three-dimensional structures of Cdk and cyclin proteins have revealed that, in the absence of cyclin, the active site in the Cdk protein is partly obscured by a slab of protein, like a stone blocking the entrance to a cave (**Figure 17–17**A). Cyclin binding causes the slab to move away from the active site, resulting in partial activation of the Cdk enzyme (Figure 17–17B). Full activation of the cyclin–Cdk complex then occurs when a separate kinase, the **Cdk-activating kinase (CAK)**, phosphorylates an amino acid near the entrance of the Cdk active site. This causes a small conformational change that further increases the activity of the Cdk, allowing the kinase to phosphorylate its target proteins effectively and thereby induce specific cell-cycle events (Figure 17–17C). <TAGA>

Inhibitory Phosphorylation and Cdk Inhibitory Proteins (CKIs) Can Suppress Cdk Activity

The rise and fall of cyclin levels is the primary determinant of Cdk activity during the cell cycle. Several additional mechanisms, however, fine-tune Cdk activity at specific stages of the cycle.

Phosphorylation at a pair of amino acids in the roof of the kinase active site inhibits the activity of a cyclin–Cdk complex. Phosphorylation of these sites by a protein kinase known as **Wee1** inhibits Cdk activity, while dephosphorylation of these sites by a phosphatase known as **Cdc25** increases Cdk activity (**Figure 17–18**). We will see later that this regulatory mechanism is particularly important in the control of M-Cdk activity at the onset of mitosis.

Binding of **Cdk inhibitor proteins (CKIs)** also regulates cyclin–Cdk complexes. The three-dimensional structure of a cyclin–Cdk–CKI complex reveals

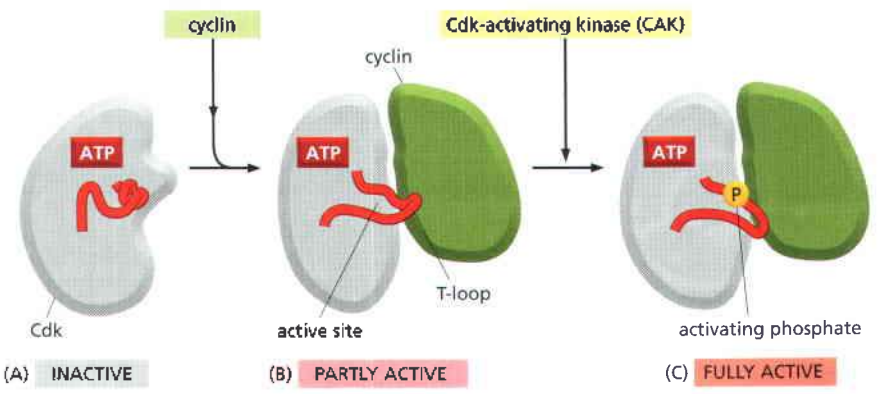


Figure 17–17 **The structural basis of Cdk activation.** These drawings are based on three-dimensional structures of human Cdk2, as determined by x-ray crystallography. The location of the bound ATP is indicated. The enzyme is shown in three states. (A) In the inactive state, without cyclin bound, the active site is blocked by a region of the protein called the T-loop *(red)*. (B) The binding of cyclin causes the T-loop to move out of the active site, resulting in partial activation of the Cdk2. (C) Phosphorylation of Cdk2 (by CAK) at a threonine residue in the T-loop further activates the enzyme by changing the shape of the T-loop, improving the ability of the enzyme to bind its protein substrates.

that CKI binding stimulates a large rearrangement in the structure of the Cdk active site, rendering it inactive (Figure 17–19). Cells use CKIs primarily to help govern the activities of G_1/S- and S-Cdks early in the cell cycle.

The Cell-Cycle Control System Depends on Cyclical Proteolysis

Whereas activation of specific cyclin-Cdk complexes drives progression through the Start and G_2/M checkpoints (see Figure 17–16), progression through the metaphase-to-anaphase transition is triggered not by protein phosphorylation but by protein destruction, leading to the final stages of cell division.

The key regulator of the metaphase-to-anaphase transition is the **anaphase-promoting complex, or cyclosome (APC/C)**, a member of the ubiquitin ligase family of enzymes. As discussed in Chapter 3, many of these enzymes are used in numerous cell processes to stimulate the proteolytic destruction of specific regulatory proteins. They transfer multiple copies of the small protein ubiquitin to specific target proteins, resulting in their proteolytic destruction by the proteasomes. Other ubiquitin ligases mark proteins for purposes other than destruction.

The APC/C catalyzes the ubiquitylation and destruction of two major proteins. The first is *securin*, which normally protects the protein linkages that hold sister chromatid pairs together in early mitosis. Destruction of securin at the metaphase-to-anaphase transition activates a protease that separates the sisters and unleashes anaphase. The S- and M-cyclins are the second major targets of the APC/C. Destroying these cyclins inactivates most Cdks in the cell (see Figure 17–16). As a result, the many proteins phosphorylated by Cdks from S phase to early mitosis are dephosphorylated by various phosphatases that are present in the anaphase cell. This dephosphorylation of Cdk targets is required for the completion of M phase, including the final steps in mitosis and the process of cytokinesis. Following its activation in mid-mitosis, the APC/C remains active in G_1, thereby providing a stable period of Cdk inactivity. When G_1/S-Cdks are activated in late G_1, the APC/C is turned off, thereby allowing cyclin accumulation to start the next cell cycle.

The cell-cycle control system also uses another ubiquitin ligase called **SCF** (after the names of its three subunits). It ubiquitylates certain CKI proteins in late G_1 and thereby helps control the activation of S-Cdks and DNA replication.

The APC/C and SCF are both large, multisubunit complexes with some related components, but they are regulated differently. APC/C activity changes during the cell cycle, primarily as a result of changes in its association with an activating subunit—either **Cdc20** during anaphase or **Cdh1** from late mitosis through early G_1. These subunits help the APC/C recognize its target proteins (Figure 17–20A). SCF activity also depends on subunits called F-box proteins, which help the complex recognize its target proteins. Unlike APC/C activity, however, SCF activity is constant during the cell cycle. Ubiquitylation by SCF is controlled instead by changes in the phosphorylation state of its target proteins, as F-box subunits recognize only specifically phosphorylated proteins (Figure 17–20B).

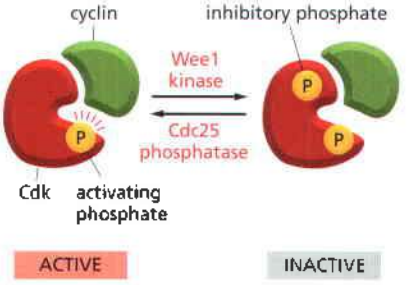


Figure 17–18 The regulation of Cdk activity by inhibitory phosphorylation. The active cyclin–Cdk complex is turned off when the kinase Wee1 phosphorylates two closely spaced sites above the active site. Removal of these phosphates by the phosphatase Cdc25 activates the cyclin–Cdk complex. For simplicity, only one inhibitory phosphate is shown. CAK adds the activating phosphate, as shown in Figure 17–17.

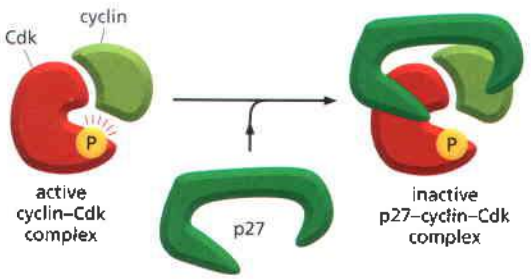


Figure 17–19 The inhibition of a cyclin–Cdk complex by a CKI. This drawing is based on the three-dimensional structure of the human cyclin A–Cdk2 complex bound to the CKI p27, as determined by x-ray crystallography. The p27 binds to both the cyclin and Cdk in the complex, distorting the active site of the Cdk. It also inserts into the ATP-binding site, further inhibiting the enzyme activity.

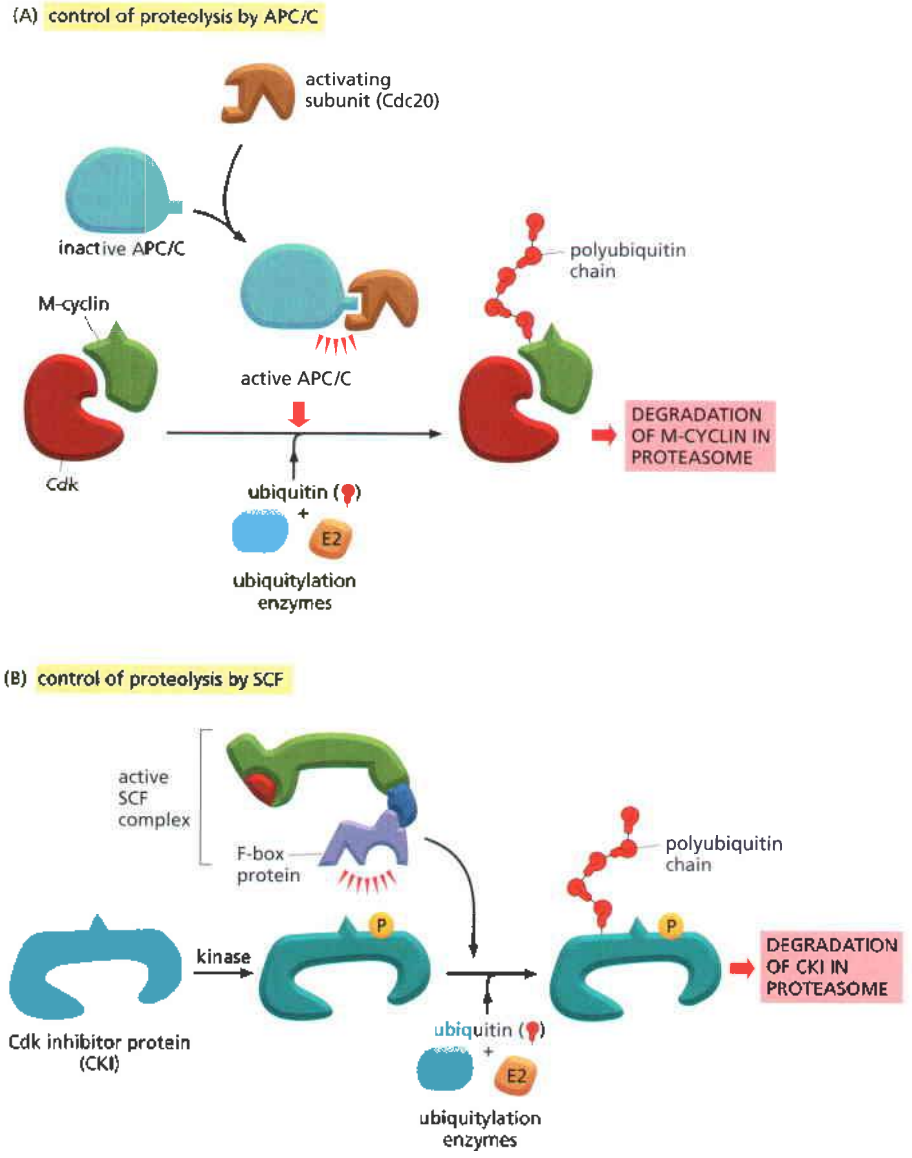


Figure 17–20 The control of proteolysis by APC/C and SCF during the cell cycle. (A) The APC/C is activated in mitosis by association with the activating subunit Cdc20, which recognizes specific amino acid sequences on M-cyclin and other target proteins. With the help of two additional proteins called E1 and E2, the APC/C transfers multiple ubiquitin molecules onto the target protein. The polyubiquitylated target is then recognized and degraded in a proteasome. (B) The activity of the ubiquitin ligase SCF depends on substrate-binding subunits called F-box proteins, of which there are many different types. The phosphorylation of a target protein, such as the CKI shown, allows the target to be recognized by a specific F-box subunit.

Cell-Cycle Control Also Depends on Transcriptional Regulation

In the frog embryonic cell cycle discussed earlier, gene transcription does not occur. Cell-cycle control depends exclusively on post-transcriptional mechanisms that involve the regulation of Cdks and ubiquitin ligases and their target proteins. In the more complex cell cycles of most cell types, however, transcriptional control provides an additional level of regulation. Changes in cyclin gene transcription, for example, help control cyclin levels in most cells.

We can use DNA microarrays (discussed in Chapter 8) to analyze changes in the expression of all of the genes in the genome as the cell progresses through the cell cycle. The results of these studies are surprising. In budding yeast, for example, about 10% of the genes encode mRNAs whose levels oscillate during the cell cycle. Some of these genes encode proteins with known cell-cycle functions, but the functions of many others are unknown.

The Cell-Cycle Control System Functions as a Network of Biochemical Switches

Table 17–2 summarizes some of the major components of the cell-cycle control system. These proteins are functionally linked together to form a robust network, which operates essentially autonomously to activate a series of biochemical switches, each of which triggers a specific cell-cycle event.

Table 17–2 **Summary of the Major Cell-Cycle Regulatory Proteins**

GENERAL NAME	FUNCTIONS AND COMMENTS
Protein kinases and protein phosphatases that modify Cdks	
Cdk-activating kinase (CAK)	phosphorylates an activating site in Cdks
Wee1 kinase	phosphorylates inhibitory sites in Cdks; primarily involved in suppressing Cdk1 activity before mitosis
Cdc25 phosphatase	removes inhibitory phosphates from Cdks; three family members (Cdc25A, B, C) in mammals; primarily involved in controlling Cdk1 activation at the onset of mitosis
Cdk inhibitor proteins (CKIs)	
Sic1 (budding yeast)	suppresses Cdk1 activity in G_1; phosphorylation by Cdk1 at the end of G_1 triggers its destruction
p27 (mammals)	suppresses G_1/S-Cdk and S-Cdk activities in G_1; helps cells withdraw from cell cycle when they terminally differentiate; phosphorylation by Cdk2 triggers its ubiquitylation by SCF
p21 (mammals)	suppresses G_1/S-Cdk and S-Cdk activities following DNA damage
p16 (mammals)	suppresses G_1-Cdk activity in G_1; frequently inactivated in cancer
Ubiquitin ligases and their activators	
APC/C	catalyzes ubiquitylation of regulatory proteins involved primarily in exit from mitosis, including securin and S- and M-cyclins; regulated by association with activating subunits
Cdc20	APC/C-activating subunit in all cells; triggers initial activation of APC/C at metaphase-to-anaphase transition; stimulated by M-Cdk activity
Cdh1	APC/C-activating subunit that maintains APC/C activity after anaphase and throughout G_1; inhibited by Cdk activity
SCF	catalyzes ubiquitylation of regulatory proteins involved in G_1 control, including some CKIs (Sic1 in budding yeast, p27 in mammals); phosphorylation of target protein usually required for this activity

When conditions for cell proliferation are right, various external and internal signals stimulate the activation of G_1-Cdk, which in turn stimulates the expression of genes encoding G_1/S- and S-cyclins. The resulting activation of G_1/S-Cdk then drives progression through the Start checkpoint. By mechanisms we discuss later, G_1/S-Cdks unleash a wave of S-Cdk activity, which initiates chromosome duplication in S phase and also contributes to some early events of mitosis. M-Cdk activation then triggers progression through the G_2/M checkpoint and the events of early mitosis, leading to the alignment of sister chromatids at the equator of the mitotic spindle. Finally, the APC/C, together with its activator Cdc20, triggers the destruction of securin and cyclins at the metaphase-to-anaphase transition, thereby unleashing sister-chromatid segregation and the completion of mitosis (**Figure 17–21**). When mitosis is complete, multiple mechanisms collaborate to suppress Cdk activity after mitosis, resulting in a stable G_1 period, as we discuss later. We are now ready to discuss these cell-cycle stages in more detail, starting with S phase.

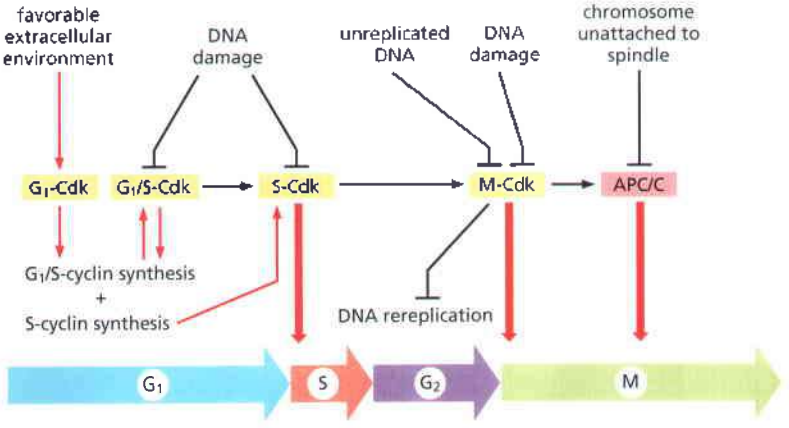


Figure 17–21 **An overview of the cell-cycle control system.** The core of the cell-cycle control system consists of a series of cyclin–Cdk complexes *(yellow)*. As discussed in more detail later, the activity of each complex is also influenced by various inhibitory mechanisms, which provide information about the extracellular environment, cell damage, and incomplete cell-cycle events *(top)*. These mechanisms are not present in all cell types; many are missing in early embryonic cell cycles, for example.

Summary

The cell-cycle control system triggers the events of the cell cycle and ensures that these events are properly timed and occur in the correct order. The control system responds to various intracellular and extracellular signals and arrests the cycle when the cell either fails to complete an essential cell-cycle process or encounters unfavorable environmental or intracellular conditions.

Central components of the cell-cycle control system are cyclin-dependent protein kinases (Cdks), which depend on cyclin subunits for their activity. Oscillations in the activities of various cyclin–Cdk complexes control various cell-cycle events. Thus, activation of S-phase cyclin–Cdk complexes (S-Cdk) initiates S phase, while activation of M-phase cyclin–Cdk complexes (M-Cdk) triggers mitosis. The mechanisms that control the activities of cyclin–Cdk complexes include phosphorylation of the Cdk subunit, binding of Cdk inhibitor proteins (CKIs), proteolysis of cyclins, and changes in the transcription of genes encoding Cdk regulators. The cell-cycle control system also depends crucially on two additional enzyme complexes, the APC/C and SCF ubiquitin ligases, which catalyze the ubiquitylation and consequent destruction of specific regulatory proteins that control critical events in the cycle.

S PHASE

The linear chromosomes of eucaryotic cells are vast and dynamic assemblies of DNA and protein, and their duplication is a complex process that takes up a major fraction of the cell cycle. Not only must the long DNA molecule of each chromosome be duplicated accurately—a remarkable feat in itself—but the protein packaging surrounding each region of that DNA must also be reproduced, ensuring that the daughter cells inherit all features of chromosome structure.

The central event of chromosome duplication is replication of the DNA. A cell must solve two problems when initiating and completing DNA replication. First, replication must occur with extreme accuracy to minimize the risk of mutations in the next cell generation. Second, every nucleotide in the genome must be copied once, and only once, to prevent the damaging effects of gene amplification. In Chapter 5, we discuss the sophisticated protein machinery that performs DNA replication with astonishing speed and accuracy. In this section, we consider the elegant mechanisms by which the cell-cycle control system initiates the replication process and, at the same time, prevents it from happening more than once per cycle.

S-Cdk Initiates DNA Replication Once Per Cycle

DNA replication begins at *origins of replication*, which are scattered at numerous locations in every chromosome. During S phase, the *initiation* of DNA replication occurs at these origins when specialized protein machines (sometimes called *initiator proteins*) unwind the double helix at the origin and load DNA replication enzymes onto the two single-stranded templates. This leads to the *elongation* phase of replication, when the replication machinery moves outward from the origin at two *replication forks* (discussed in Chapter 5).

To ensure that chromosome duplication occurs only once per cell cycle, the initiation phase of DNA replication is divided into two distinct steps that occur at different times in the cell cycle. The first step occurs in late mitosis and early G_1, when a large complex of initiator proteins, called the **prereplicative complex**, or **pre-RC**, assembles at origins of replication. This step is sometimes called *licensing* of replication origins because initiation of DNA synthesis is permitted only at origins containing a pre-RC. The second step occurs at the onset of S phase, when components of the pre-RC nucleate the formation of a larger protein complex called the **preinitiation complex**. This complex then unwinds the DNA helix and loads DNA polymerases and other replication enzymes onto the DNA strands, thereby initiating DNA synthesis, as described in Chapter 5. Once the replication origin has been activated in this way, the

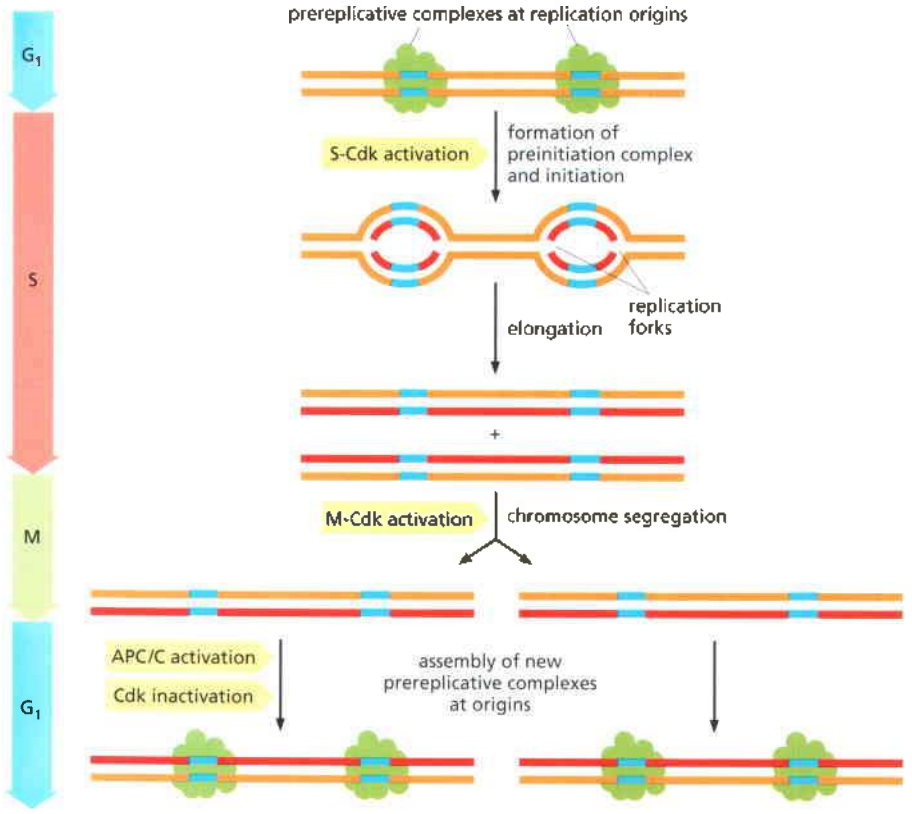


Figure 17–22 Control of chromosome duplication. Preparations for DNA replication begin in G_1 with the assembly of prereplicative complexes (pre-RCs) at replication origins. S-Cdk activation leads to the formation of multiprotein preinitiation complexes that unwind the DNA at origins and begin the process of DNA replication. Two replication forks move out from each origin until the entire chromosome is duplicated. Duplicated chromosomes are then segregated in M phase. The activation of replication origins in S phase also causes disassembly of the prereplicative complex, which does not reform at the origin until the following G_1—thereby ensuring that each origin is activated only once in each cell cycle.

pre-RC is dismantled and cannot be reassembled at that origin until the following G_1. As a result, origins can be activated only once per cell cycle.

The cell-cycle control system governs both assembly of the pre-RC and assembly of the pre-initiation complex (Figure 17–22). Assembly of the pre-RC is inhibited by Cdk activity, and, in most cells, is stimulated by the APC/C. Pre-RC assembly therefore occurs only in late mitosis and early G_1, when Cdk activity is low and APC/C activity is high. At the onset of S phase, activation of S-Cdk then triggers the formation of a preinitiation complex, which initiates DNA synthesis. In addition, the pre-RC is partly dismantled. Because S-Cdk and M-Cdk activities remain high (and APC/C activity remains low) until late mitosis, new pre-RCs cannot be assembled at fired origins until the cell cycle is complete.

Figure 17–23 illustrates some of the proteins involved in the initiation of DNA replication. A key player is a large, multiprotein complex called the **origin recognition complex (ORC)**, which binds to replication origins throughout the cell cycle. In late mitosis and early G_1, the proteins **Cdc6** and **Cdt1** bind to the ORC at origins and help load a group of six related proteins called the **Mcm proteins**. The resulting large complex is the pre-RC, and the origin is now licensed for replication.

The six Mcm proteins of the pre-RC form a ring around the DNA that is thought to serve as the major DNA helicase that unwinds the origin DNA when DNA synthesis begins and as the replication forks move out from the origin. Thus, the central purpose of the pre-RC is to load the helicase that will play a central part in the subsequent DNA replication process.

Once the pre-RC has assembled in G_1, the replication origin is ready to fire. The activation of S-Cdk in late G_1 triggers the assembly of several additional protein complexes at the origin, leading to the formation of a giant preinitiation complex that unwinds the helix and begins DNA synthesis.

At the same time as it initiates DNA replication, S-Cdk triggers the disassembly of some pre-RC components at the origin. Cdks phosphorylate both the ORC and Cdc6, resulting in their inhibition by various mechanisms. Furthermore, inactivation of the APC/C in late G_1 also helps turn off pre-RC assembly. In late mitosis and early G_1, the APC/C triggers the destruction of a protein,

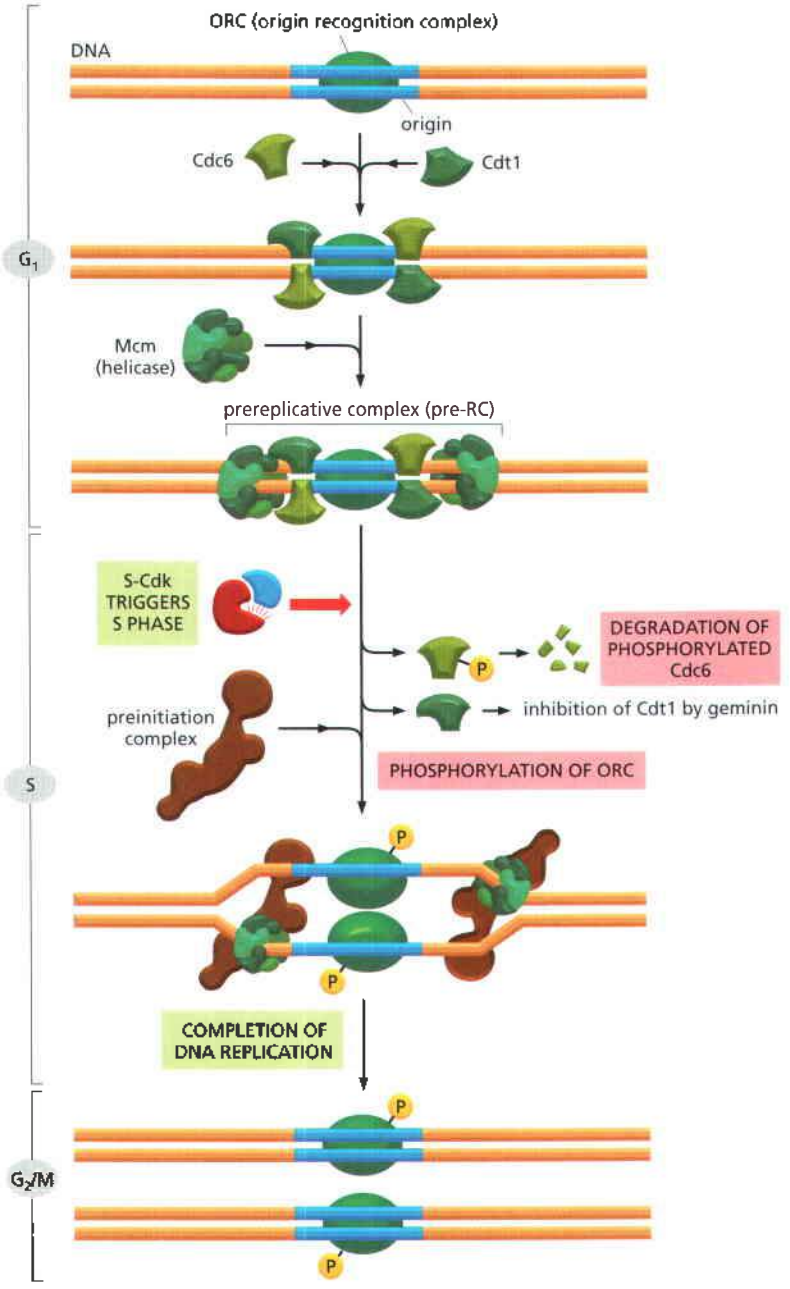


Figure 17–23 Control of the initiation of DNA replication. The ORC remains associated with a replication origin throughout the cell cycle. In early G_1, Cdc6 and Cdt1 associate with the ORC. The resulting protein complex then assembles Mcm ring complexes on the adjacent DNA, resulting in the formation of the prereplicative complex (pre-RC). S-Cdk (with assistance from another protein kinase, not shown) then stimulates the assembly of several additional proteins at the origin to form the preinitiation complex. DNA polymerase and other replication proteins are recruited to the origin, the Mcm protein rings are activated as DNA helicases, and DNA unwinding allows DNA replication to begin. S-Cdk also blocks rereplication by triggering the destruction of Cdc6 and the inactivation of the ORC. Cdt1 is inactivated by the protein geminin. Geminin is an APC/C target and its levels therefore increase in S and M phases, when APC/C is inactive. Thus, the components of the pre-RC (Cdc6, Cdt1, Mcm) cannot form a new pre-RC at the origins until M-Cdk is inactivated and the APC/C is activated at the end of mitosis (see text).

geminin, that binds and inhibits the pre-RC component Cdt1. Thus, when the APC/C is turned off in late G_1, geminin accumulates and inhibits Cdt1. In these various ways, S- and M-Cdk activities, combined with low APC/C activity, block pre-RC formation during S phase and thereafter. How, then, is the cell-cycle control system reset to allow replication to occur in the next cell cycle? The answer is simple. At the end of mitosis, APC/C activation leads to the inactivation of Cdks and the destruction of geminin. Pre-RC components are dephosphorylated and Cdt1 is activated, allowing pre-RC assembly to prepare the cell for the next S phase.

Chromosome Duplication Requires Duplication of Chromatin Structure

The DNA of the chromosomes is extensively packaged in a variety of protein components, including histones and various regulatory proteins involved in the control of gene expression (discussed in Chapter 4). Thus, duplication of a

chromosome is not simply a matter of duplicating the DNA at its core but also requires the duplication of these chromatin proteins and their proper assembly on the DNA.

The production of chromatin proteins increases during S phase to provide the raw materials needed to package the newly synthesized DNA. Most importantly, S-Cdks stimulate a large increase in the synthesis of the four histone subunits that form the histone octamers at the core of each nucleosome. These subunits are assembled into nucleosomes on the DNA by nucleosome assembly factors, which typically associate with the replication fork and distribute nucleosomes on both strands of the DNA as they emerge from the DNA synthesis machinery.

Chromatin packaging helps to control gene expression. In some parts of the chromosome, the chromatin is highly condensed and is called *heterochromatin*, whereas in other regions it has a more open structure and is called *euchromatin*. These differences in chromatin structure depend on a variety of mechanisms, including modification of histone tails and the presence of non-histone proteins (discussed in Chapter 4). Because these differences are important in gene regulation, it is crucial that chromatin structure, like the DNA within, is reproduced accurately during S phase. How chromatin structure is duplicated is not well understood, however. During DNA synthesis, histone-modifying enzymes and various non-histone proteins are probably deposited onto the two new DNA strands as they emerge from the replication fork, and these proteins are thought to help reproduce the local chromatin structure of the parent chromosome.

Cohesins Help Hold Sister Chromatids Together

At the end of S phase, each replicated chromosome consists of a pair of identical sister chromatids glued together along their length. This sister-chromatid cohesion sets the stage for a successful mitosis because it greatly facilitates the attachment of the two sister chromatids in a pair to opposite poles of the mitotic spindle. Imagine how difficult it would be to achieve this bipolar attachment if sister chromatids were allowed to drift apart after S phase. Indeed, defects in sister-chromatid cohesion—in yeast mutants, for example—lead inevitably to major errors in chromosome segregation.

Sister-chromatid cohesion depends on a large protein complex called **cohesin**, which is deposited at many locations along the length of each sister chromatid as the DNA is replicated in S phase. Two of the subunits of cohesin are members of a large family of proteins called *SMC proteins* (for Structural Maintenance of Chromosomes). Cohesin forms giant ring-like structures, and it has been proposed that these might form rings that surround the two sister chromatids (Figure 17–24).

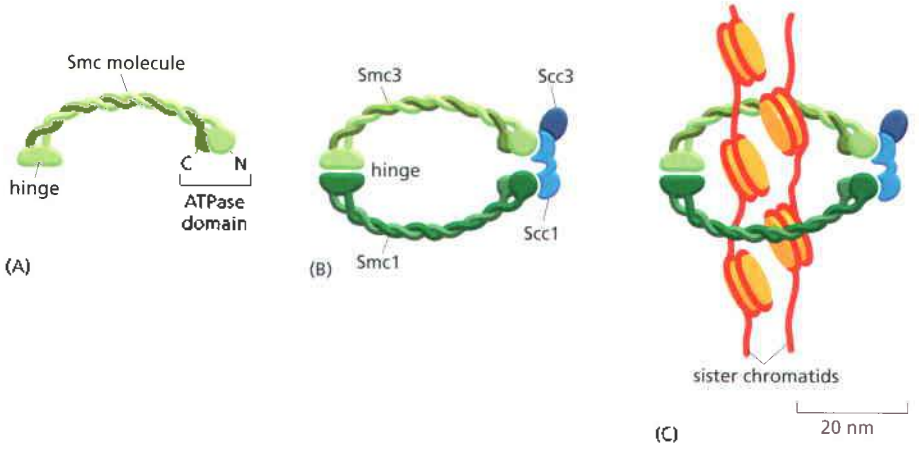


Figure 17–24 **Cohesin.** Cohesin is a protein complex with four subunits. Two subunits, Smc1 and Smc3, are coiled-coil proteins with an ATPase domain at one end; together, they form a large V-shaped structure as shown. Two additional subunits, Scc1 and Scc3, connect the ATPase head domains forming a ring structure that may encircle the sister chromatids as shown.

Sister-chromatid cohesion also results, at least in part, from *DNA catenation*, the intertwining of sister DNA molecules that occurs when two replication forks meet during DNA synthesis. The enzyme topoisomerase II gradually disentangles the catenated sister DNAs between S phase and early mitosis by cutting one DNA molecule, passing the other through the break, and then resealing the cut DNA (see Figure 5–23). Once the catenation has been removed, sister-chromatid cohesion depends primarily on cohesin complexes. The loss of sister cohesion at the metaphase-to-anaphase transition therefore depends primarily on disruption of these complexes, as we describe later.

Summary

Duplication of the chromosomes in S phase involves the accurate copying of the entire DNA molecule in each chromosome, as well as the duplication of the chromatin proteins that associate with the DNA and govern various aspects of chromosome function. Chromosome duplication is triggered by the activation of S-Cdk, which activates proteins that unwind the DNA and initiate its replication at sites in the DNA called replication origins. Once a replication origin is activated during S phase, S-Cdk also inhibits proteins that are required to allow that origin to initiate DNA replication again. Thus, each origin is fired once and only once in each S phase and cannot be reused until the next cell cycle.

MITOSIS

Following the completion of S phase and transition through G_2, the cell undergoes the dramatic upheaval of M phase. This begins with mitosis, during which the sister chromatids are separated and distributed (*segregated*) to a pair of identical daughter nuclei, each with its own copy of the genome. Mitosis is traditionally divided into five stages—*prophase, prometaphase, metaphase, anaphase,* and *telophase*—defined primarily on the basis of chromosome behavior as seen in a microscope. <TACT> <TCAA> As mitosis is completed, the second major event of M phase—cytokinesis—divides the cell into two halves, each with an identical nucleus. **Panel 17–1** summarizes the major events of M phase.

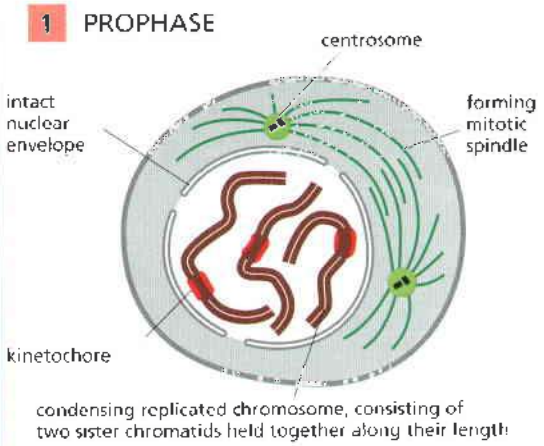
From a regulatory point of view, mitosis can be divided into two major parts, each governed by distinct components of the cell-cycle control system. First, an abrupt increase in M-Cdk activity at the G_2/M checkpoint triggers the events of early mitosis (prophase, prometaphase, and metaphase). During this period, M-Cdk and several other mitotic protein kinases phosphorylate a variety of proteins, leading to the assembly of the mitotic spindle and its attachment to the sister chromatid pairs. The second major part of mitosis begins at the metaphase-to-anaphase transition, when the APC/C triggers the destruction of securin, liberating a protease that cleaves cohesin and thereby initiates separation of the sister chromatids. The APC/C also triggers the destruction of cyclins, which leads to Cdk inactivation and the dephosphorylation of Cdk targets, which is required for all events of late M phase, including the completion of anaphase, the disassembly of the mitotic spindle, and the division of the cell by cytokinesis.

In this section, we describe the key mechanical events of mitosis and how M-Cdk and the APC/C orchestrate them.

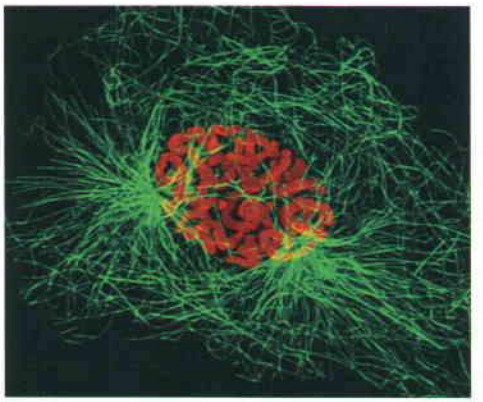
M-Cdk Drives Entry Into Mitosis

One of the most remarkable features of cell-cycle control is that a single protein kinase, M-Cdk, brings about all of the diverse and complex cell rearrangements that occur in the early stages of mitosis. At a minimum, M-Cdk must induce the assembly of the mitotic spindle and ensure that each sister chromatid in a pair is attached to the opposite pole of the spindle. It also triggers *chromosome condensation*, the large-scale reorganization of the intertwined sister chromatids into compact, rod-like structures. In animal cells, M-Cdk also promotes the breakdown of the nuclear envelope and rearrangements of

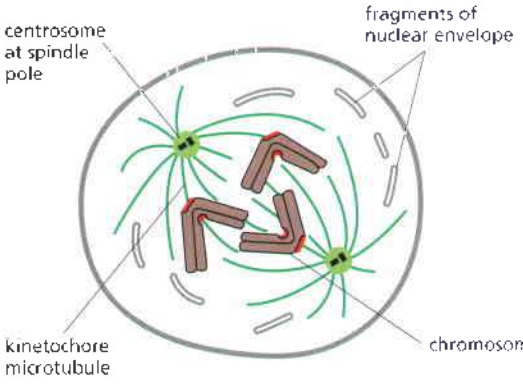
1 PROPHASE



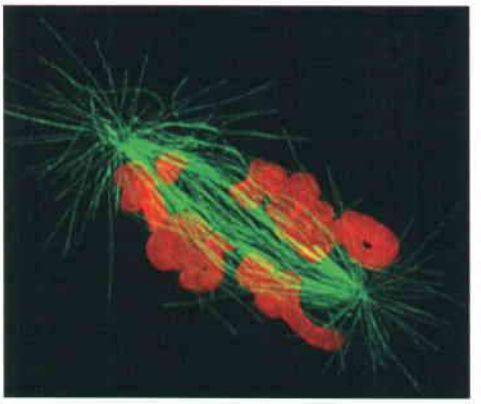
At prophase, the replicated chromosomes, each consisting of two closely associated sister chromatids, condense. Outside the nucleus, the mitotic spindle assembles between the two centrosomes, which have replicated and moved apart. For simplicity, only three chromosomes are shown. In diploid cells, there would be two copies of each chromosome present. In the photomicrograph, chromosomes are stained *orange* and microtubules are *green*.



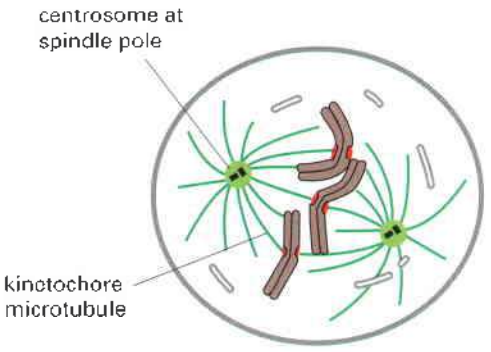
2 PROMETAPHASE



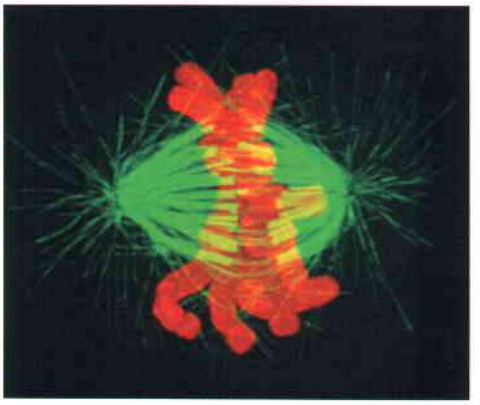
Prometaphase starts abruptly with the breakdown of the nuclear envelope. Chromosomes can now attach to spindle microtubules via their kinetochores and undergo active movement.



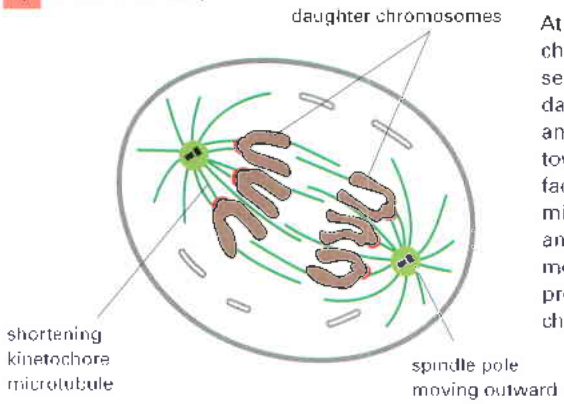
3 METAPHASE



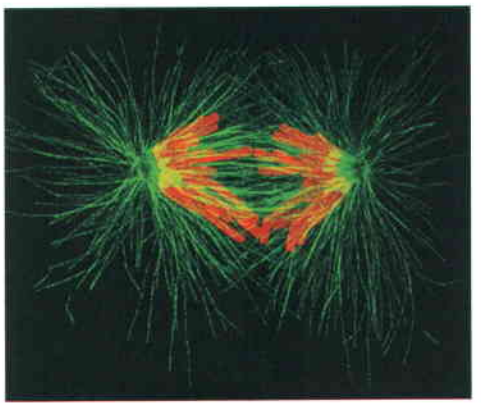
At metaphase, the chromosomes are aligned at the equator of the spindle, midway between the spindle poles. The kinetochore microtubules attach sister chromatids to opposite poles of the spindle.



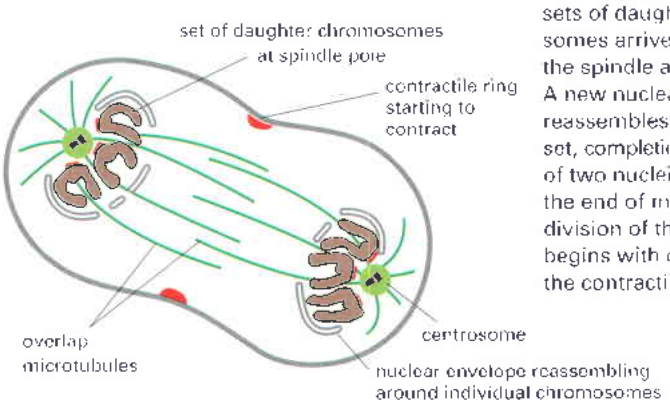
4 ANAPHASE



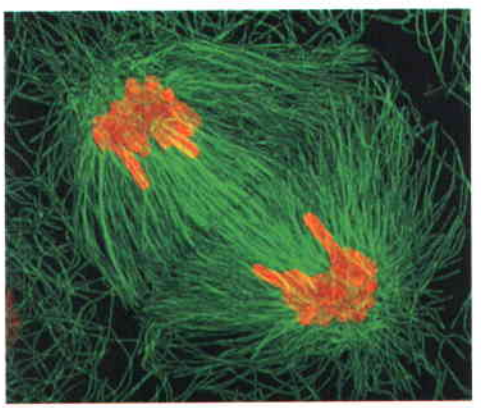
At anaphase, the sister chromatids synchronously separate to form two daughter chromosomes, and each is pulled slowly toward the spindle pole it faces. The kinetochore microtubules get shorter, and the spindle poles also move apart; both processes contribute to chromosome segregation.



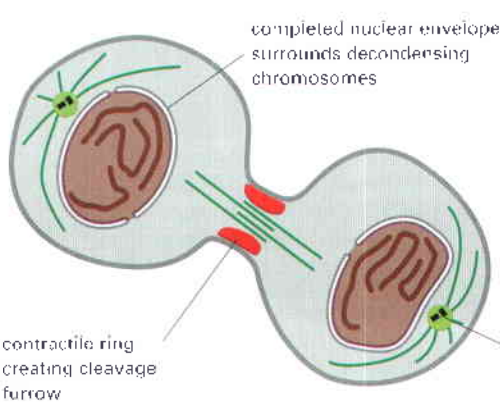
5 TELOPHASE



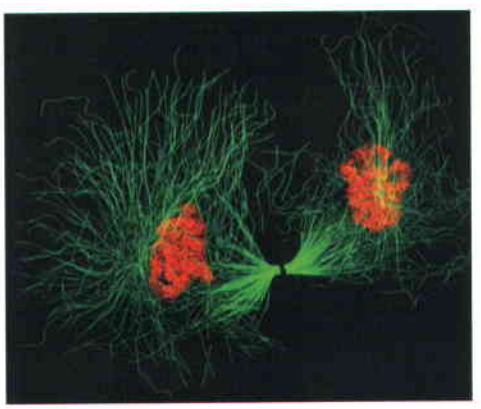
During telophase, the two sets of daughter chromosomes arrive at the poles of the spindle and decondense. A new nuclear envelope reassembles around each set, completing the formation of two nuclei and marking the end of mitosis. The division of the cytoplasm begins with contraction of the contractile ring.



6 CYTOKINESIS



During cytokinesis, the cytoplasm is divided in two by a contractile ring of actin and myosin filaments, which pinches the cell in two to create two daughters, each with one nucleus.



(Micrographs courtesy of Julie Canman and Ted Salmon.)

the actin cytoskeleton and the Golgi apparatus. Each of these processes is thought to be triggered when M-Cdk phosphorylates specific proteins involved in the process, although most of these proteins have not yet been identified.

M-Cdk does not act alone to phosphorylate key proteins involved in early mitosis. Two additional families of protein kinases, the *Polo-like kinases* and the *Aurora kinases*, also make important contributions to the control of early mitotic events. The Polo-like kinase Plk, for example, is required for the normal assembly of a bipolar mitotic spindle, in part because it phosphorylates proteins involved in separation of the spindle poles early in mitosis. The Aurora kinase Aurora-A also helps control proteins that govern the assembly and stability of the spindle, whereas Aurora-B controls attachment of sister chromatids to the spindle, as we discuss later. Activation of Polo-like kinases and Aurora kinases depends on M-Cdk activity, but the precise activation mechanisms are not clear.

Dephosphorylation Activates M-Cdk at the Onset of Mitosis

M-Cdk activation begins with the accumulation of M-cyclin (cyclin B in vertebrate cells; see Table 17–1). In embryonic cell cycles, the synthesis of M-cyclin is constant throughout the cell cycle, and M-cyclin accumulation results from the high stability of the protein in interphase. In most cell types, however, M-cyclin synthesis increases during G_2 and M, owing primarily to an increase in *M-cyclin* gene transcription. The increase in M-cyclin protein leads to a corresponding accumulation of M-Cdk (the complex of Cdk1 and M-cyclin) as the cell approaches mitosis. Although the Cdk in these complexes is phosphorylated at an activating site by the Cdk-activating kinase (CAK), as discussed earlier, the protein kinase Wee1 holds it in an inactive state by inhibitory phosphorylation at two neighboring sites (see Figure 17–18). Thus, by the time the cell reaches the end of G_2, it contains an abundant stockpile of M-Cdk that is primed and ready to act but is suppressed by phosphates that block the active site of the kinase.

What, then, triggers the activation of the M-Cdk stockpile? The crucial event is the activation of the protein phosphatase Cdc25, which removes the inhibitory phosphates that restrain M-Cdk (**Figure 17–25**). At the same time, the inhibitory activity of the kinase Wee1 is suppressed, further ensuring that M-Cdk activity increases. The mechanisms that unleash Cdc25 activity (and suppress Wee1) in early mitosis are not well understood. One possibility is that the S-Cdks that are active in G_2 and early prophase stimulate Cdc25.

Interestingly, Cdc25 can also be activated, at least in part, by its target, M-Cdk. M-Cdk may also inhibit the inhibitory kinase Wee1. The ability of M-Cdk to activate its own activator (Cdc25) and inhibit its own inhibitor (Wee1) suggests that M-Cdk activation in mitosis involves positive feedback loops (see Figure 17–25). According to this attractive model, the partial activation of Cdc25 (perhaps by S-Cdk) leads to the partial activation of a subpopulation of M-Cdk

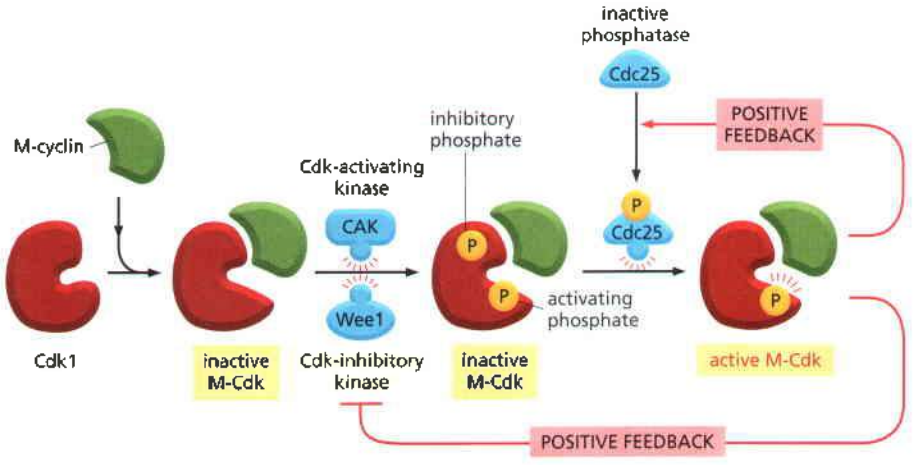


Figure 17–25 The activation of M-Cdk. Cdk1 associates with M-cyclin as the levels of M-cyclin gradually rise. The resulting M-Cdk complex is phosphorylated on an activating site by the Cdk-activating kinase (CAK) and on a pair of inhibitory sites by the Wee1 kinase. The resulting inactive M-Cdk complex is then activated at the end of G_2 by the phosphatase Cdc25. Cdc25 is further stimulated by active M-Cdk, resulting in positive feedback. This feedback is enhanced by the ability of M-Cdk to inhibit Wee1.

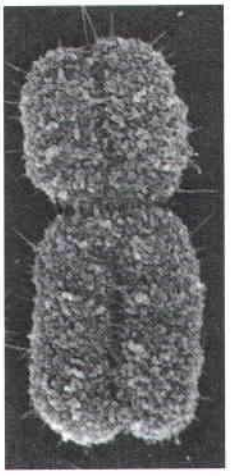


Figure 17–26 The mitotic chromosome. Scanning electron micrograph of a human mitotic chromosome, consisting of two sister chromatids joined along their length. The constricted regions are the centromeres. (Courtesy of Terry D. Allen.)

complexes, which then phosphorylate more Cdc25 and Wee1 molecules. This leads to more M-Cdk activation, and so on. Such a mechanism would quickly promote the complete activation of all the M-Cdk complexes in the cell. As mentioned earlier, similar molecular switches operate at various points in the cell cycle to promote the abrupt and complete transition from one cell-cycle state to the next.

Condensin Helps Configure Duplicated Chromosomes for Separation

At the end of S phase, the immensely long DNA molecules of the sister chromatids are tangled in a mass of partially catenated DNA and proteins. Any attempt to pull the sisters apart in this state would undoubtedly lead to breaks in the chromosomes. To avoid this disaster, the cell devotes a great deal of energy in early mitosis to gradually reorganizing the sister chromatids into relatively short, distinct structures that can be pulled apart more easily in anaphase. These chromosomal changes involve two processes: *chromosome condensation*, in which the chromatids are dramatically compacted; and *sister-chromatid resolution*, whereby the two sisters are resolved into distinct, separable units (**Figure 17–26**). Resolution results from the decatenation of the sister DNAs, accompanied by the partial removal of cohesin molecules along the chromosome arms. As a result, when the cell reaches metaphase, the sister chromatids appear in the microscope as compact, rod-like structures that are joined tightly at their centromeric regions and only loosely along their arms.

The condensation and resolution of sister chromatids depends, at least in part, on a five-subunit protein complex called **condensin**. Condensin structure is related to that of the cohesin complex that holds sister chromatids together (see Figure 17–24). It contains two SMC subunits like those of cohesin, plus three non-SMC subunits (**Figure 17–27**). Condensin may form a ring-like structure that somehow uses the energy provided by ATP hydrolysis to promote the compaction and resolution of sister chromatids. Condensin is able to change the coiling of DNA molecules in a test tube, and this coiling activity is thought to be important for chromosome condensation during mitosis. Interestingly, phosphorylation of condensin subunits by M-Cdk stimulates this coiling activity, providing one mechanism by which M-Cdk may promote chromosome restructuring in early mitosis.

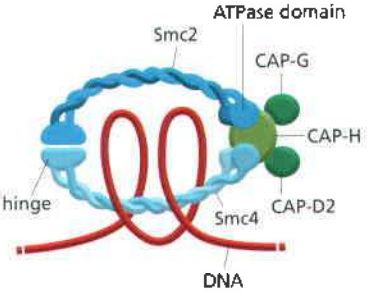


Figure 17–27 Condensin. Condensin is a five-subunit protein complex that resembles cohesin (see Figure 17–24). The head domains of its two major subunits, Smc2 and Smc4, are held together by three additional subunits. It is not clear how condensin catalyzes the restructuring and compaction of chromosome DNA, but it may form a ring structure that encircles loops of DNA as shown; it can hydrolyze ATP and coil DNA molecules in a test tube.

The Mitotic Spindle Is a Microtubule-Based Machine

The central event of mitosis—chromosome segregation—depends in all eucaryotes on a complex and beautiful machine called the **mitotic spindle**. The spindle is a bipolar array of microtubules, which pulls sister chromatids apart in anaphase, thereby segregating the two sets of chromosomes to opposite ends of the cell, where they are packaged into daughter nuclei. M-Cdk triggers the assembly of the spindle early in mitosis, in parallel with the chromosome restructuring just described. Before we consider how the spindle assembles and how its microtubules attach to sister chromatids, we briefly review the basic features of spindle structure.

As discussed in Chapter 16, the core of the mitotic spindle is a bipolar array of microtubules, the minus ends of which are focused at the two spindle poles, and the plus ends of which radiate outward from the poles (**Figure 17–28**). <GTCT> The plus ends of some microtubules—called the **interpolar microtubules**—interact with the plus ends of microtubules from the other pole, resulting in an antiparallel array in the spindle midzone. The plus ends of other

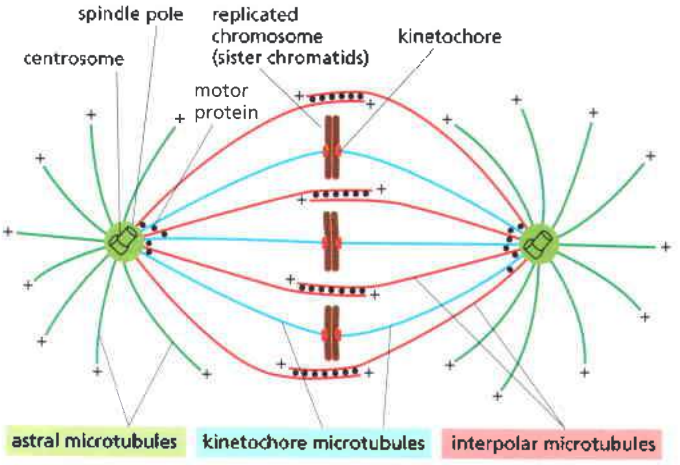


Figure 17–28 The three classes of microtubules of the mitotic spindle in an animal cell. The plus ends of the microtubules project away from the centrosomes, while the minus ends are anchored at the spindle poles, which in this example are organized by centrosomes. Kinetochore microtubules connect the spindle poles with the kinetochores of sister chromatids, while interpolar microtubules from the two poles interdigitate at the spindle equator. Astral microtubules radiate out from the poles into the cytoplasm and usually interact with the cell cortex, helping to position the spindle in the cell.

microtubules—the **kinetochore microtubules**—are attached to sister chromatid pairs at large protein structures called *kinetochores*, which are located at the *centromere* of each sister chromatid. Finally, many spindles also contain **astral microtubules** that radiate outward from the poles and contact the cell cortex, helping to position the spindle in the cell.

In most somatic animal cells, each spindle pole is focused at a protein organelle called the **centrosome** (discussed in Chapter 16). Each centrosome consists of a cloud of amorphous material (called the *pericentriolar matrix)* that surrounds a pair of *centrioles* (**Figure 17–29**). The pericentriolar matrix nucleates a radial array of microtubules, with their fast-growing plus ends projecting outward and their minus ends associated with the centrosome. The matrix contains a variety of proteins, including microtubule-dependent motor proteins,

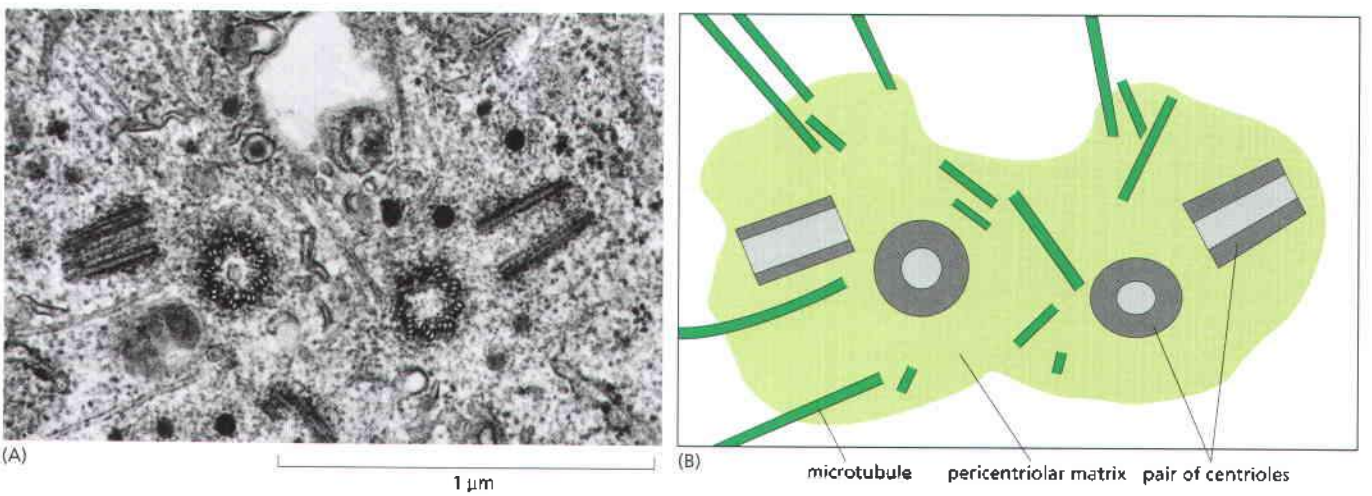
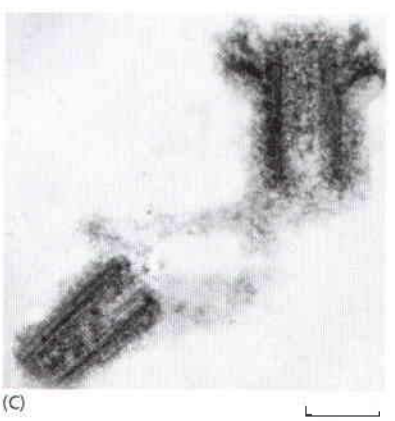


Figure 17–29 The centrosome. (A) Electron micrograph of an S-phase mammalian cell in culture, showing a duplicated centrosome. Each centrosome contains a pair of centrioles; although the centrioles have duplicated, they remain together in a single complex, as shown in the drawing of the micrograph in (B). One centriole of each centriole pair has been cut in cross section, while the other is cut in longitudinal section, indicating that the two members of each pair are aligned at right angles to each other. The two halves of the replicated centrosome, each consisting of a centriole pair surrounded by pericentriolar matrix, will split and migrate apart to initiate the formation of the two poles of the mitotic spindle when the cell enters M phase. (C) Electron micrograph of a centriole pair that has been isolated from a cell. The two centrioles have partly separated during the isolation procedure but remain tethered together by fine fibers, which keep the centriole pair together until it is time for them to separate. Both centrioles are cut longitudinally, and it can now be seen that the two have different structures: the mother centriole is larger and more complex than the daughter centriole, and only the mother centriole is associated with pericentriolar matrix that nucleates microtubules. Each daughter centriole will mature during the next cell cycle, when it will replicate to give rise to its own daughter centriole. (A, from M. McGill, D.P. Highfield, T.M. Monahan and B.R. Brinkley, *J. Ultrastruct. Res.* 57:43–53, 1976. With permission from Academic Press; C, from M. Paintrand et al., *J. Struct. Biol.* 108:107–128, 1992. With permission from Elsevier.)



coiled-coil proteins that link the motors to the centrosome, structural proteins, and components of the cell-cycle control system. Most important, it contains the *γ-tubulin ring complex*, which is the component mainly responsible for nucleating microtubules (discussed in Chapter 16).

Some cells—notably the cells of higher plants and the oocytes of many vertebrates—do not have centrosomes, and microtubule-dependent motor proteins and other proteins associated with microtubule minus ends organize and focus the spindle poles.

Microtubule-Dependent Motor Proteins Govern Spindle Assembly and Function

The assembly and function of the mitotic spindle depend on numerous microtubule-dependent motor proteins. As discussed in Chapter 16, these proteins belong to two families—the kinesin-related proteins, which usually move toward the plus end of microtubules, and dyneins, which move toward the minus end. In the mitotic spindle, these motor proteins generally operate at or near the ends of the microtubules. Four major types of motor proteins—*kinesin-5*, *kinesin-14*, *kinesins-4 and 10*, and *dynein*—are particularly important in spindle assembly and function (**Figure 17–30**).

Kinesin-5 proteins contain two motor domains that interact with the plus ends of antiparallel microtubules in the spindle midzone. Because the two motor domains move toward the plus ends of the microtubules, they slide the two antiparallel microtubules past each other toward the spindle poles, forcing the poles apart. Kinesin-14 proteins, by contrast, are minus-end directed motors with a single motor domain and other domains that can interact with a different microtubule. They can cross-link antiparallel interpolar microtubules at the spindle midzone and tend to pull the poles together. Kinesin-4 and kinesin-10 proteins, also called *chromokinesins*, are plus-end directed motors that associate with chromosome arms and push the attached chromosome away from the pole (or the pole away from the chromosome). Finally, dyneins are minus-end directed motors that, together with associated proteins, organize microtubules at various cellular locations. They link the plus ends of astral microtubules to components of the actin cytoskeleton at the cell cortex, for example; by moving toward the minus end of the microtubules, the dynein motors pull the spindle poles toward the cell cortex and away from each other.

Two Mechanisms Collaborate in the Assembly of a Bipolar Mitotic Spindle

The mitotic spindle must have two poles if it is to pull the two sets of sister chromatids to opposite ends of the cell in anaphase. In animal cells, the primary focus of this chapter, two mechanisms collaborate to ensure the bipolarity of the spindle. One depends on the ability of mitotic chromosomes to nucleate and stabilize microtubules and on the ability of the various motor proteins just described to organize microtubules into a bipolar array, with minus ends

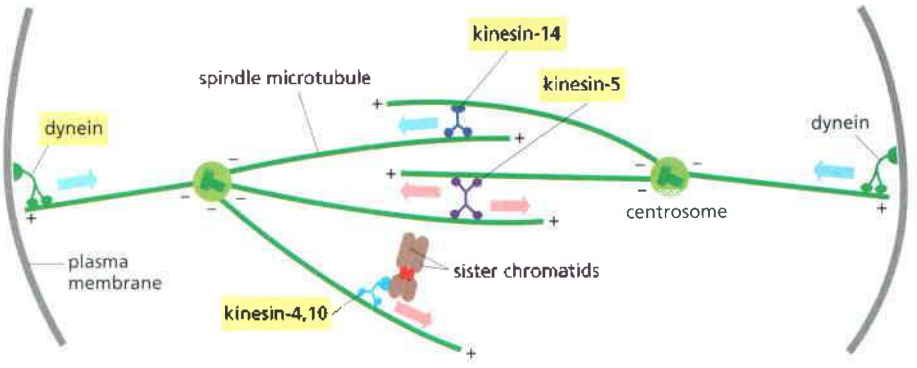


Figure 17–30 Major motor proteins of the spindle. Four major classes of microtubule-dependent motor proteins (*yellow* boxes) contribute to spindle assembly and function (see text). The colored arrows indicate the direction of motor movement along a microtubule—*blue* toward the minus end, and *red* toward the plus end.

focused at two spindle poles and plus ends interacting with each other in the spindle midzone. The other depends on the ability of centrosomes to help form the spindle poles. A typical animal cell enters mitosis with a pair of centrosomes, each of which nucleates a radial array of microtubules. The two centrosomes facilitate bipolar spindle assembly by providing a pair of prefabricated spindle poles. Centrosomes are not essential for the assembly of a bipolar spindle, however, since a functional spindle forms in cells that normally lack centrosomes and in cultured cells in which a laser beam has destroyed the centrosome.

We now describe the steps of spindle assembly, beginning with centrosome-dependent assembly in early mitosis. We then consider the self-organization mechanism that does not require centrosomes and becomes particularly important after nuclear envelope breakdown.

Centrosome Duplication Occurs Early in the Cell Cycle

Most animal cells contain a single centrosome that nucleates most of the cell's cytoplasmic microtubules. The centrosome duplicates when the cell enters the cell cycle, so that by the time the cell reaches mitosis there are two centrosomes. Centrosome duplication begins at about the same time as the cell enters S phase. The G_1/S-Cdk (a complex of cyclin E and Cdk2 in animal cells; see Table 17–1) that triggers cell cycle entry also initiates centrosome duplication. The two centrioles in the centrosome separate, and each nucleates the formation of a single new centriole, resulting in two centriole pairs within an enlarged pericentriolar matrix (**Figure 17–31**). This centrosome pair remains together on one side of the nucleus until the cell enters mitosis.

There are interesting parallels between centrosome duplication and chromosome duplication. Both use a semi-conservative mechanism of duplication, in which the two halves separate and serve as templates for construction of a new half. Centrosomes, like chromosomes, must replicate once and only once per cell cycle, to ensure that the cell enters mitosis with only two copies: an incorrect number of centrosomes could lead to defects in spindle assembly and thus errors in chromosome segregation.

The mechanisms that limit centrosome duplication to once per cell cycle are uncertain. In many cell types, experimental inhibition of DNA synthesis blocks centrosome duplication, providing one mechanism by which centrosome number is kept in check. Other cell types, however, including those in the early embryos of flies, sea urchins, and frogs, do not contain such a mechanism and centrosome duplication continues if chromosome duplication is blocked. It is not known how such cells limit centrosome duplication to once per cell cycle.

M-Cdk Initiates Spindle Assembly in Prophase

At the beginning of mitosis, the sudden rise in M-Cdk activity initiates spindle assembly. In animal cells, the two centrosomes move apart along the nuclear

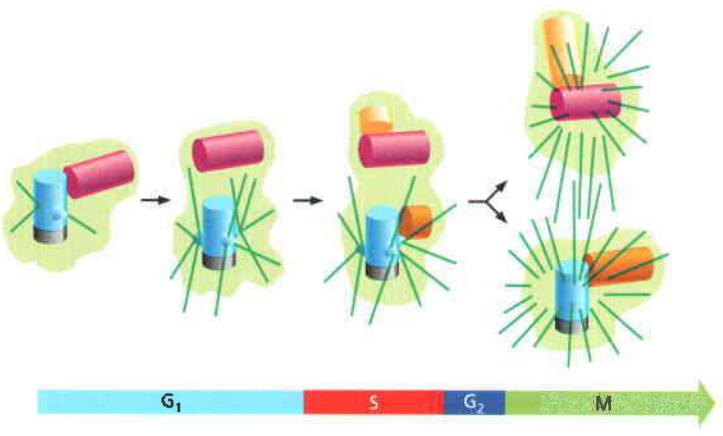


Figure 17–31 Centriole replication. The centrosome consists of a centriole pair and associated pericentriolar matrix *(green)*. At a certain point in G_1, the two centrioles of the pair separate by a few micrometers. During S phase, a daughter centriole begins to grow near the base of each mother centriole and at a right angle to it. The elongation of the daughter centriole is usually completed by G_2. The two centriole pairs remain close together in a single centrosomal complex until the beginning of M phase, when the complex splits in two and the two halves begin to separate. Each centrosome now nucleates its own radial array of microtubules called an aster.

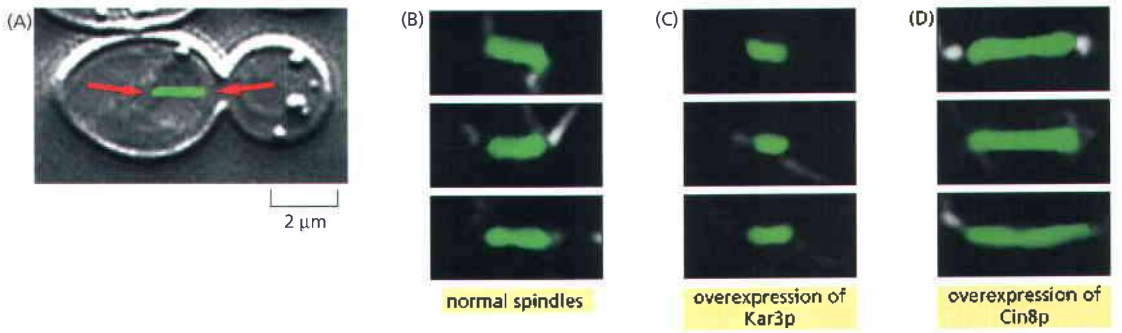


Figure 17–32 The influence of opposing motor proteins on spindle length in budding yeast. (A) A differential-interference-contrast micrograph of a mitotic yeast cell. The spindle is highlighted in *green*, and the position of the spindle poles is indicated by *red arrows*. The nuclear envelope does not break down during mitosis in yeasts, and the spindle forms inside the nucleus. In (B–D), the mitotic spindles have been stained with fluorescent anti-tubulin antibodies. (B) Normal yeast cells. (C) Overexpression of the minus-end-directed motor protein Kar3 (a kinesin-14 protein) leads to abnormally short spindles. (D) Overexpression of the plus-end-directed motor protein Cin8 (a kinesin-5 protein) leads to abnormally long spindles. Thus, it seems that a balance between opposing motor proteins determines spindle length in these cells. (A, courtesy of Kerry Bloom; B–D, from W. Saunders, V. Lengyel and M.A. Hoyt, *Mol. Biol. Cell* 8:1025–1033, 1997. With permission from American Society for Cell Biology.)

envelope, and the plus ends of the microtubules between them interdigitate to form the interpolar microtubules of the developing spindle. At the same time, the amount of γ-tubulin ring complexes in each centrosome increases greatly, increasing the ability of the centrosomes to nucleate new microtubules, a process called *centrosome maturation*.

Multiple motor proteins drive the separation of centrosomes in early mitosis. In prophase, minus-end directed dynein motor proteins at the plus ends of astral microtubules provide the major force. These motors are anchored at the cell cortex or on the nuclear envelope, and their movement toward the microtubule minus end pulls the centrosomes apart (see Figure 17–30). Following nuclear envelope breakdown at the end of prophase, interactions between the centrosomal microtubules and the cell cortex allow actin–myosin bundles in the cortex to pull the centrosomes further apart. Finally, kinesin-5 motors cross-link the overlapping, antiparallel ends of interpolar microtubules and push the poles apart (see Figure 17–30).

The balance of opposing forces generated by different types of motor proteins determines the final length of the spindle. Dynein and kinesin-5 motors generally promote centrosome separation and increase spindle length. Kinesin-14 proteins do the opposite: they are minus-end directed motors and interact with a microtubule from one pole while traveling toward the minus end of an antiparallel microtubule from the other pole; as a result, they tend to pull the poles together. It is not clear how the cell regulates the balance of opposing forces to generate the appropriate spindle length (**Figure 17–32**).

M-Cdk and other mitotic protein kinases are required for centrosome separation and maturation. M-Cdk and aurora-A phosphorylate kinesin-5 motors and stimulate them to drive centrosome separation. Aurora-A and Plk also phosphorylate components of the centrosome and thereby promote its maturation.

The Completion of Spindle Assembly in Animal Cells Requires Nuclear Envelope Breakdown

The centrosomes and microtubules of animal cells are located in the cytoplasm, separated from the chromosomes by the double membrane barrier of the nuclear envelope (discussed in Chapter 12). Clearly, the attachment of sister chromatids to the spindle requires the removal of this barrier. In addition, many of the motor proteins and microtubule regulators that promote spindle assembly are associated with the chromosomes inside the nucleus. Nuclear envelope breakdown allows these proteins to carry out their important functions in spindle assembly.

Nuclear envelope breakdown is a complex, multi-step process that is thought to begin when M-Cdk phosphorylates several subunits of the giant nuclear pore complexes in the nuclear envelope. This initiates the disassembly of nuclear pore complexes and their dissociation from the envelope. M-Cdk also phosphorylates components of the nuclear lamina, the structural framework that lies beneath the envelope. The phosphorylation of these lamina components and of several inner nuclear envelope proteins leads to disassembly of the nuclear lamina and the breakdown of the envelope membranes into small vesicles.

Microtubule Instability Increases Greatly in Mitosis

Most animal cells in interphase contain a cytoplasmic array of microtubules radiating out from the single centrosome. As discussed in Chapter 16, the microtubules of this interphase array are in a state of *dynamic instability*, in which individual microtubules are either growing or shrinking and stochastically switch between the two states. The switch from growth to shrinkage is called a *catastrophe*, and the switch from shrinkage to growth is called a *rescue* (see Figure 16–16). New microtubules are continually being created to balance the loss of those that disappear completely by depolymerization.

Entry into mitosis signals an abrupt change in the cell's microtubules. The interphase array of few, long microtubules is converted to a larger number of shorter and more dynamic microtubules surrounding each centrosome. During prophase, and particularly in prometaphase and metaphase (see Panel 17–1), the half-life of microtubules decreases dramatically. This increase in microtubule instability, coupled with the increased ability of centrosomes to nucleate microtubules as mentioned earlier, results in remarkably dense and dynamic arrays of spindle microtubules that are ideally suited for capturing sister chromatids.

M-Cdk initiates these changes in microtubule behavior, at least in part, by phosphorylating two classes of proteins that control microtubule dynamics (discussed in Chapter 16). These include microtubule-dependent motor proteins and **microtubule-associated proteins (MAPs)**. Experiments using cell-free *Xenopus* egg extracts, which reproduce many of the changes that occur in intact cells during the cell cycle, have revealed the roles of these regulators in controlling microtubule dynamics. If centrosomes and fluorescent tubulin are added to these extracts, fluorescent microtubules nucleate from the centrosomes, and we can observe the behavior of individual microtubules by time-lapse fluorescence video microscopy. The microtubules in mitotic extracts differ from those in interphase extracts primarily by the increased rate of catastrophes, in which the microtubules switch abruptly from slow growth to rapid shortening.

Two classes of proteins govern microtubule dynamics in mitosis. Proteins called **catastrophe factors** destabilize microtubule arrays by increasing the frequency of catastrophes (see Figure 16–16). One of these proteins is a kinesin-related protein that does not function as a motor. MAPs, by contrast, have the opposite effect, stabilizing microtubules in various ways: they can increase the frequency of rescues, in which microtubules switch from shrinkage to growth, or they can either increase the growth rate or decrease the shrinkage rate of microtubules. Thus, in principle, changes in catastrophe factors and MAPs can make microtubules much more dynamic in M phase by increasing total microtubule depolymerization rates, decreasing total microtubule polymerization rates, or both.

In *Xenopus* egg extracts, the balance between a single type of catastrophe factor and a single type of MAP determines the catastrophe rate and the steady-state length of microtubules (**Figure 17–33**). This balance, in turn, governs the assembly of the mitotic spindle, as microtubules that are either too long or too short cannot form a normal spindle. One way in which M-Cdk may control microtubule length is by phosphorylating this MAP and reducing its ability to stabilize microtubules. Even if the activity of the catastrophe factor remained constant throughout the cell cycle, the balance between the two opposing activities of the MAP and catastrophe factor would shift, increasing the dynamic instability of the microtubules.

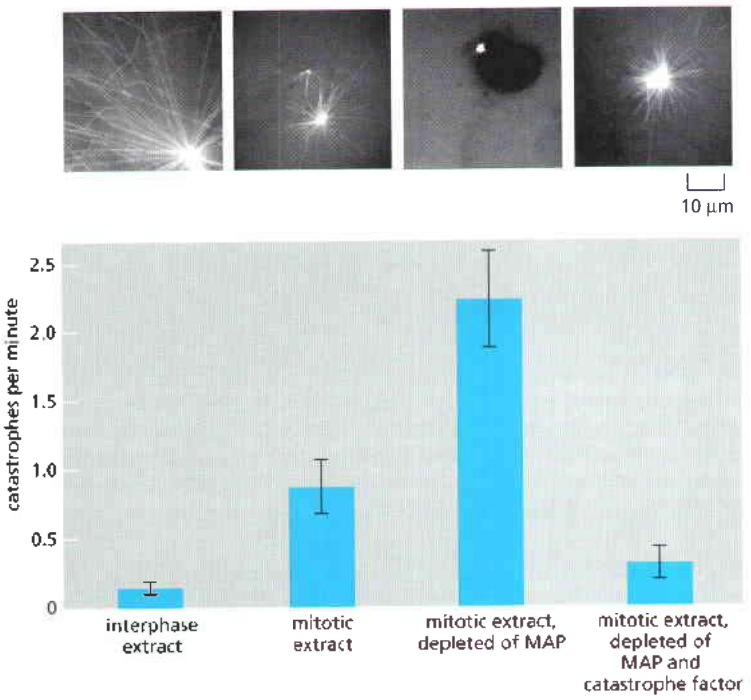


Figure 17–33 Experimental evidence that the balance between catastrophe factors and MAPs influences the frequency of microtubule catastrophes and microtubule length. Interphase or mitotic *Xenopus* egg extracts were incubated with centrosomes and fluorescent tubulin, and the behavior of individual microtubules nucleated from the centrosomes was followed by fluorescence video microscopy. As expected, the catastrophe rate is higher in mitotic than in interphase extracts. The depletion of a specific MAP (called Xmap215) from the mitotic extracts increases the catastrophe rate, indicating that this MAP inhibits catastrophes in mitotic extracts. Inhibition of a specific catastrophe factor (the kinesin-related protein Mcak) greatly reduces the catastrophe rate in the MAP-depleted mitotic extracts, indicating that this factor is responsible for stimulating catastrophes in mitotic extracts. Thus, the catastrophe rate depends on the balance between the MAP and the catastrophe factor. Fluorescence micrographs of the asters formed in the different experimental conditions are shown in the top panels; note that the higher the catastrophe rates, the shorter the microtubules. (From R. Tournebize et al., *Nat. Cell Biol.* 2:13–19, 2000. With permission from Macmillan Publishers Ltd.)

Mitotic Chromosomes Promote Bipolar Spindle Assembly

Chromosomes are not just passive passengers in the process of spindle assembly. By creating a local environment that favors both microtubule nucleation and microtubule stabilization, they play an active part in spindle formation. The influence of the chromosomes can be demonstrated by using a fine glass needle to reposition them after the spindle has formed. For some cells in metaphase, if a single chromosome is tugged out of alignment, a mass of new spindle microtubules rapidly appears around the newly positioned chromosome, while the spindle microtubules at the chromosome's former position depolymerize. This property of the chromosomes seems to depend, at least in part, on a guanine-nucleotide exchange factor (GEF) that is bound to chromatin; the GEF stimulates a small GTPase in the cytosol called *Ran* to bind GTP in place of GDP. The activated Ran–GTP, which is also involved in nuclear transport (discussed in Chapter 12), releases microtubule-stabilizing proteins from protein complexes in the cytosol, thereby stimulating the local nucleation and stabilization of microtubules around chromosomes.

It is this ability of chromosomes to stabilize and organize microtubules that enables cells to form bipolar spindles in the absence of centrosomes, as discussed earlier. Acentrosomal spindle assembly is thought to begin with the nucleation and stabilization of microtubules around the chromosomes. Motor proteins, particularly members of the kinesin-5 family (see Figure 17–30), then cross-link microtubules in an antiparallel orientation and push their minus ends apart. Kinesins-4 and 10 on the chromosome arms also help push the minus ends away from the chromosomes. Dynein, kinesin-14, and various minus-end binding proteins then cross-link and focus the minus ends of the microtubules to form the two spindle poles (**Figure 17–34**).

Cells that normally lack centrosomes, such as those of higher plants and many animal oocytes, use this chromosome-based self-organization process to form spindles. It is also the process used to assemble spindles in certain insect embryos that have been induced to develop from eggs without fertilization (that is, *parthenogenetically*); as the sperm normally provides the centrosome when it fertilizes an egg (discussed in Chapter 21), the mitotic spindles in these parthenogenic embryos develop without centrosomes (**Figure 17–35**). Even in cells that normally contain centrosomes, the chromosomes help organize the

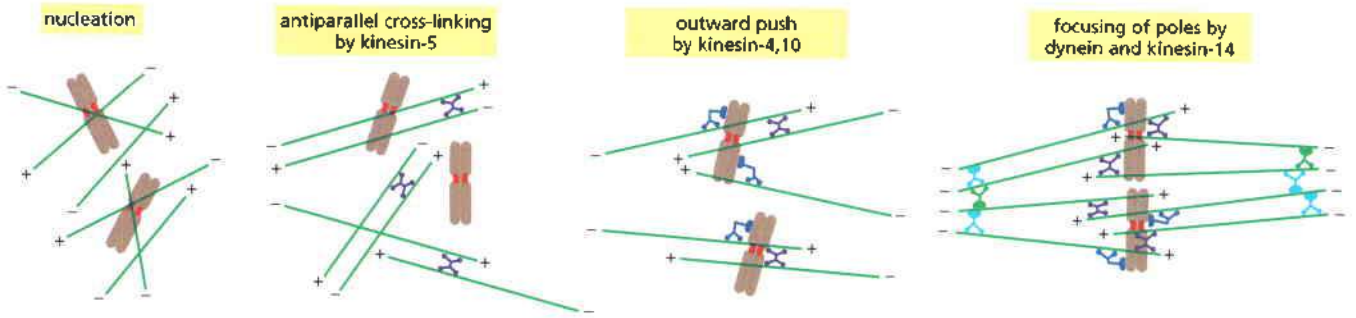


Figure 17–34 Spindle self-organization by motor proteins. Mitotic chromosomes stimulate the local production of RanGTP (not shown), which activates proteins that nucleate and promote the formation of microtubules in the vicinity of the chromosomes. Kinesin-5 motor proteins (see Figure 17–30) organize these microtubules into antiparallel bundles, while plus-end directed kinesins-4 and 10 link the microtubules to chromosome arms and push minus ends away from the chromosomes. Dynein and kinesin-14 motors, together with numerous other proteins, focus these minus ends into a pair of spindle poles.

spindle microtubules and, with the help of various motor proteins, can promote the assembly of a bipolar mitotic spindle if the centrosomes are removed. Although the resulting acentrosomal spindle can segregate chromosomes normally, it lacks astral microtubules, which are responsible for positioning the spindle in animal cells; as a result, the spindle is often mispositioned, resulting in abnormalities in cytokinesis.

Kinetochores Attach Sister Chromatids to the Spindle

Following the assembly of a bipolar microtubule array, the second major step in spindle formation is the attachment of the array to the chromosomes. Spindle microtubules are attached to each sister chromatid at the *kinetochore*, a giant, multilayered protein structure that is built on the heterochromatin that forms at the centromeric region of the chromosome (**Figure 17–36**). The plus ends of kinetochore microtubules are embedded head-on in specialized microtubule-attachment sites within the kinetochore. Animal cell kinetochores contain 10–40 of these attachment sites, whereas yeast kinetochores contain just one. Each attachment site contains a protein collar that surrounds the microtubule near its end, thereby holding the microtubule tightly to the kinetochore while still allowing the addition or removal of tubulin subunits at this end (**Figure 17–37**). Regulation of plus end polymerization and depolymerization at the kinetochore is critical for the control of chromosome movement on the spindle, as we discuss later.

Cells containing centrosomes employ a "search and capture" mechanism to attach their mitotic chromosomes to the spindle. The dynamic plus ends of microtubules radiate outward from the centrosomes and eventually capture the kinetochore of one sister chromatid. In newt lung cells, where we can observe the initial capture event in a microscope, the kinetochore is seen first to bind to the side of the microtubule and then to slide rapidly along it toward the centrosome. The lateral attachment to the chromosome is rapidly converted to an end-on attachment. At the same time, microtubules growing from

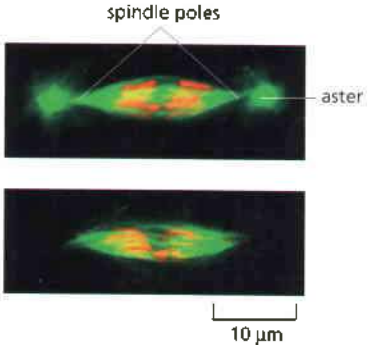


Figure 17–35 Bipolar spindle assembly without centrosomes in parthenogenetic embryos of the insect *Sciara* (or fungus gnat). The microtubules are stained *green*, the chromosomes *red*. The *top* fluorescence micrograph shows a normal spindle formed with centrosomes in a normally fertilized *Sciara* embryo. The *bottom* micrograph shows a spindle formed without centrosomes in an embryo that initiated development without fertilization. Note that the spindle with centrosomes has an aster at each pole of the spindle, whereas the spindle formed without centrosomes does not. Both types of spindles are able to segregate the replicated chromosomes. (From B. de Saint Phalle and W. Sullivan, *J. Cell Biol.* 141:1383–1391, 1998. With permission from The Rockefeller University Press.)

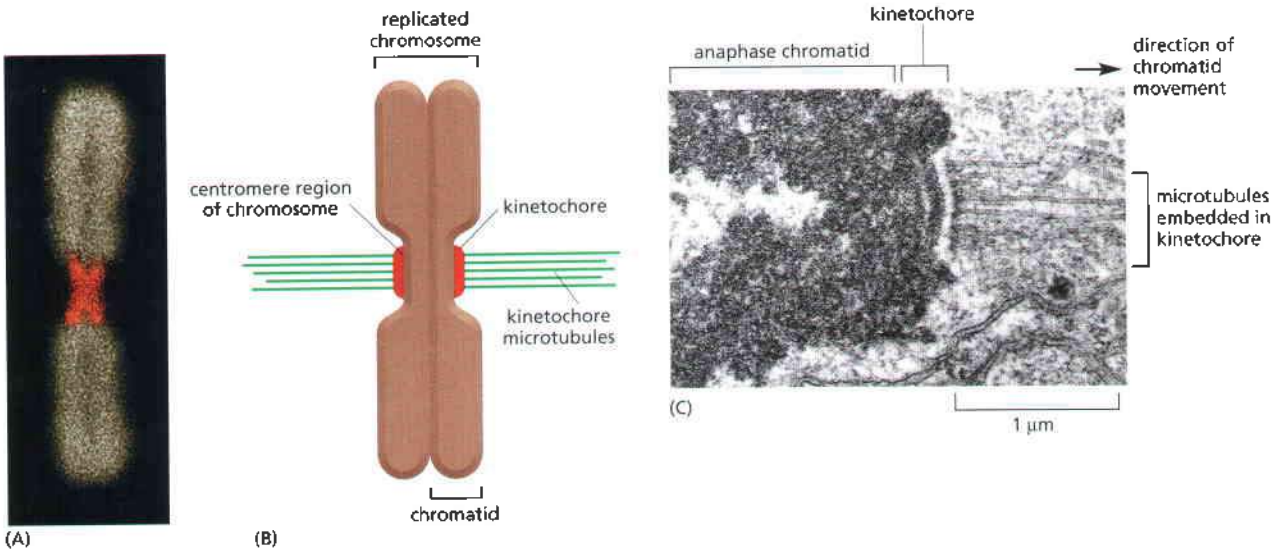


Figure 17–36 The kinetochore. (A) A fluorescence micrograph of a metaphase chromosome stained with a DNA-binding fluorescent dye and with human autoantibodies that react with specific kinetochore proteins. The two kinetochores, one associated with each chromatid, are stained *red*. (B) A drawing of a metaphase chromosome showing its two sister chromatids attached to the plus ends of kinetochore microtubules. Each kinetochore forms a plaque on the surface of the centromere. The number of microtubules bound to a metaphase kinetochore varies from 1 in budding yeast to 40 in some mammalian cells. (C) Electron micrograph of an anaphase chromatid with microtubules attached to its kinetochore. While most kinetochores have a trilaminar structure, the one shown here (from a green alga) has an unusually complex structure with additional layers. (A, courtesy of B.R. Brinkley; C, from J.D. Pickett-Heaps and L.C. Fowke, *Aust. J. Biol. Sci.* 23:71–92, 1970. With permission from CSIRO.)

the opposite spindle pole attach to the kinetochore on the opposite side of the chromosome, forming a bipolar attachment (**Figure 17–38**).

How does chromosome attachment occur in the absence of centrosomes? One possibility is that short microtubules in the vicinity of the chromosomes interact with kinetochores and become embedded in the plus-end-binding collars of the kinetochore. Polymerization at these plus ends would then result in growth of the microtubules away from the kinetochore. The minus ends of these kinetochore microtubules, like other minus ends in centrosome-free spindles, would eventually become cross-linked to other minus ends and focused by motor proteins at the spindle pole (see Figure 17–34).

Bi-Orientation Is Achieved by Trial and Error

The success of mitosis demands that sister chromatids in a pair attach to opposite poles of the mitotic spindle, so that they move to opposite ends of the cell when they separate in anaphase. How is this mode of attachment, called **bi-orientation**, achieved? What prevents the attachment of both kinetochores to the same spindle pole or the attachment of one kinetochore to both spindle poles? Part of the answer is that sister kinetochores are constructed in a back-to-back orientation that reduces the likelihood that both kinetochores can face the same spindle pole. Nevertheless, incorrect attachments do occur, and elegant regulatory mechanisms have evolved to correct them.

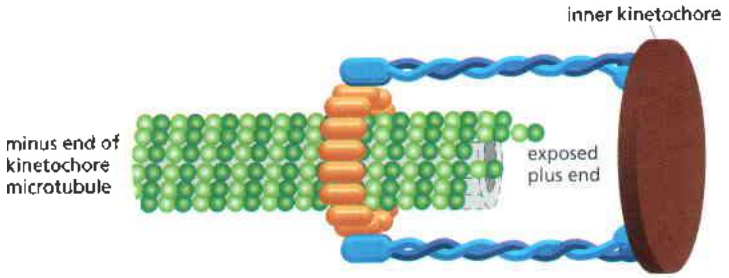


Figure 17–37 A microtubule attachment site in a kinetochore. Each site is thought to contain a collar structure *(yellow)* that surrounds the microtubule plus end, allowing polymerization and depolymerization to occur at the exposed plus end while the microtubule remains attached to the kinetochore.

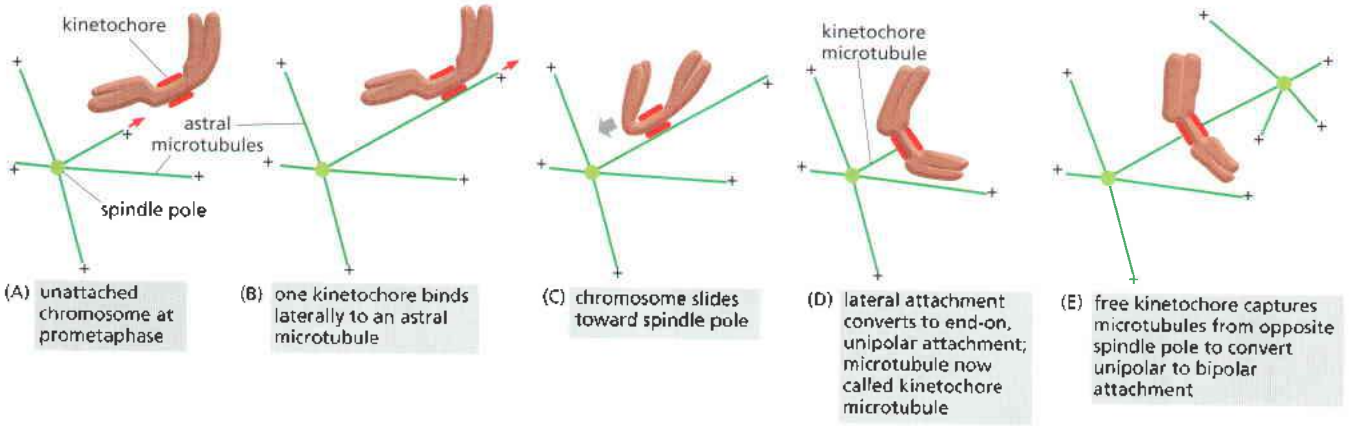


Figure 17–38 The capture of centrosome microtubules by kinetochores. The *red arrow* in (A) indicates the direction of microtubule growth, while the *gray arrow* in (C) indicates the direction of chromosome sliding.

Incorrect attachments are corrected by a system of trial and error that is based on a simple principle: incorrect attachments are highly unstable and do not last, while correct attachments are locked in place. But, how does the kinetochore sense a correct attachment? The answer appears to be tension (**Figure 17–39**). When a sister chromatid pair is properly bi-oriented on the spindle, the two kinetochores are pulled in opposite directions by strong poleward forces. Sister-chromatid cohesion resists these poleward forces, creating high levels of tension within the kinetochores. When chromosomes are incorrectly attached—when both sister chromatids are attached to the same spindle pole, for example—tension is low and the kinetochore sends an inhibitory signal that loosens the grip of its microtubule attachment site, allowing detachment to occur. When bi-orientation occurs, the high tension at the kinetochore shuts off the inhibitory signal, strengthening microtubule attachment. In animal cells, tension not only increases the affinity of the attachment site but also leads to the attachment of additional microtubules to the kinetochore. This results in the formation of a thick *kinetochore fiber* composed of multiple microtubules.

The tension-sensing mechanism depends on the protein kinase aurora-B, which is associated with the kinetochore. Aurora-B is thought to generate the inhibitory signal that reduces the strength of microtubule attachment in the absence of tension. It phosphorylates several components of the microtubule attachment site, decreasing the site's affinity for a microtubule plus end. Aurora-B is inactivated when bi-orientation occurs, thereby reducing kinetochore phosphorylation and increasing the affinity of the attachment site.

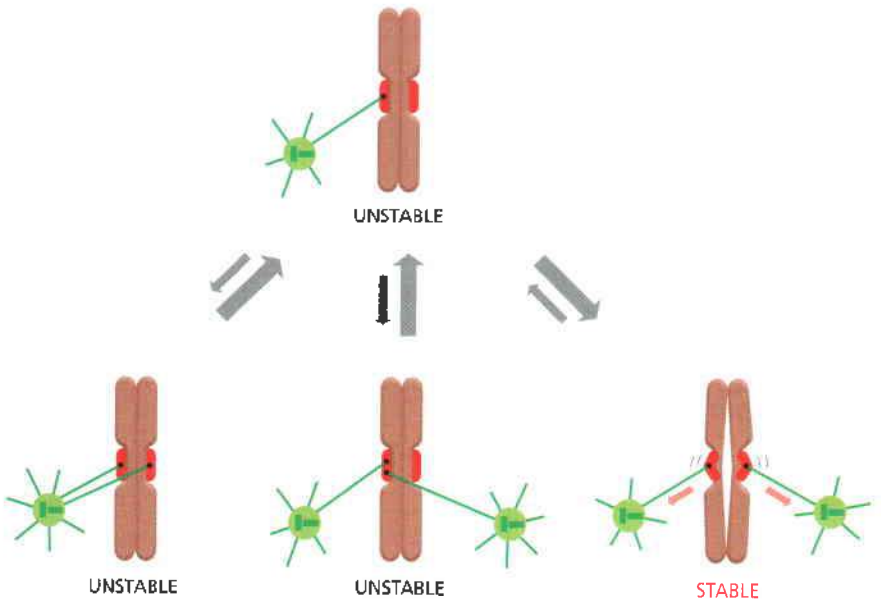


Figure 17–39 Alternative forms of chromosome attachment. Initially, a single microtubule from a spindle pole binds to one kinetochore in a sister chromatid pair. Additional microtubules can then bind to the chromosome in various ways. A microtubule from the same spindle pole can attach to the other sister kinetochore, or microtubules from both spindle poles can attach to one kinetochore. These incorrect attachments are unstable, however, so that one of the two microtubules tends to dissociate. When a second microtubule from the opposite pole binds to the second kinetochore, the sister kinetochores are thought to sense tension across their microtubule-binding sites, which triggers an increase in microtubule binding affinity. This correct attachment is thereby locked in place.

Following their attachment to the two spindle poles, the chromosomes are tugged back and forth, eventually assuming a position equidistant between the two spindle poles, a position called the **metaphase plate**. In vertebrate cells, the chromosomes then oscillate gently at the metaphase plate, awaiting the signal for the sister chromatids to separate. The signal is produced, with a predictable lag time, after the bipolar attachment of the last of the chromosomes, as we discuss later.

Multiple Forces Move Chromosomes on the Spindle

Motor proteins and other mechanisms generate the forces that move chromosomes on the microtubules of the mitotic spindle. Three major forces are thought to be particularly important.

The first major force pulls the kinetochore and its associated chromosome along the kinetochore microtubule toward the spindle pole. It is produced by proteins at the kinetochore itself. By an uncertain mechanism, depolymerization at the plus end of the microtubule somehow generates a force that pulls the kinetochore poleward (**Figure 17–40**). This force pulls on chromosomes during prometaphase and metaphase and is particularly important for moving sister chromatids toward the poles after they separate in anaphase, as we discuss later. Interestingly, this kinetochore-generated poleward force does not require ATP. This might seem implausible at first, but it has been shown that purified kinetochores in a test tube, with no ATP present, can remain attached to depolymerizing microtubules and thereby move. The energy that drives the movement is stored in the microtubule and is released when the microtubule depolymerizes; it ultimately comes from the hydrolysis of GTP that occurs after a tubulin subunit adds to the end of a microtubule (discussed in Chapter 16).

A second poleward force is provided in some cell types by **microtubule flux**, whereby the microtubules themselves are moved toward the spindle poles and dismantled at their minus ends. Until the onset of anaphase, the addition of new tubulin at the plus end of a microtubule compensates for the loss of tubulin at the minus end, so that microtubule length remains constant despite the movement of microtubules toward the spindle pole. Microtubule flux in metaphase spindles can be seen by an ingenious method in which very small amounts of fluorescent tubulin are injected into living cells (**Figure 17–41**). This results in the appearance of tiny fluorescent speckles that travel toward the poles on both kinetochore and interpolar microtubules. Any kinetochore that is attached to a microtubule undergoing such flux experiences a poleward force, which contributes to the generation of tension at the kinetochore and the poleward movement of sister chromatids after they separate in anaphase.

A third force acting on chromosomes is the *polar ejection force*. Plus-end-directed kinesin-4 and 10 motors on chromosome arms interact with interpolar microtubules and transport the chromosomes away from the spindle poles. This force is particularly important in prometaphase and metaphase, when it helps align the bi-oriented sister chromatid pairs at the metaphase plate (**Figure 17–42**).

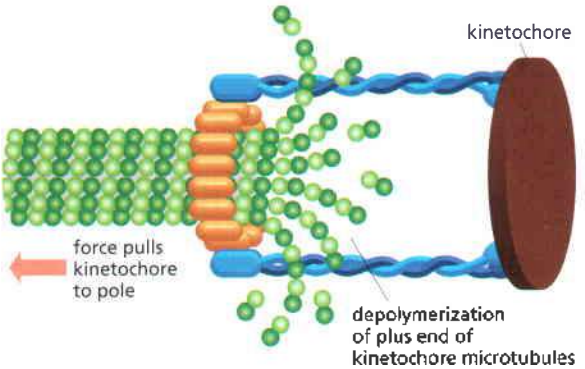
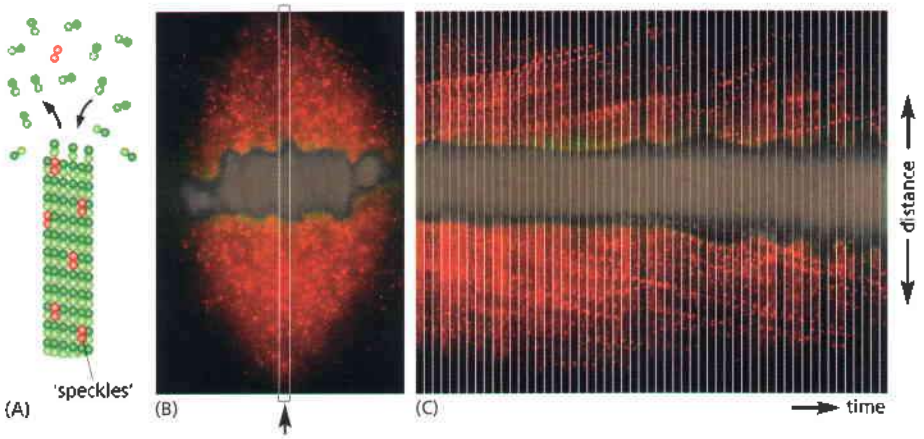


Figure 17–40 **How depolymerization may pull the kinetochore toward the spindle pole.** When depolymerization occurs, the protofilaments of the microtubule curl outward (see Figure 16–16) and push against the collar structure that surrounds the microtubule plus end. In principle, this will move the kinetochore toward the microtubule minus end at the spindle pole.



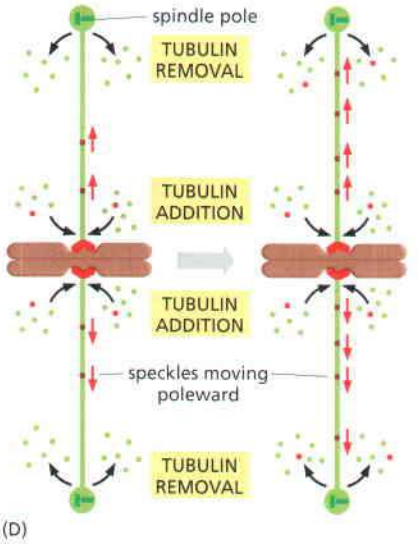


Figure 17–41 Microtubule flux in the metaphase spindle. (A) To observe microtubule flux, a very small amount of fluorescent tubulin is injected into living cells so that individual microtubules form with a very small proportion of fluorescent tubulin. Such microtubules have a speckled appearance when viewed by fluorescence microscopy. (B) Fluorescence micrographs of a mitotic spindle in a living newt lung epithelial cell. The chromosomes are colored *brown*, and the tubulin speckles are *red*. (C) The movement of individual speckles can be followed by time-lapse video microscopy. Images of the long, thin, rectangular, boxed region *(arrow)* in (B) were taken at sequential times and pasted side by side to make a montage of the region over time. Individual speckles can be seen to move toward the poles at a rate of about 0.75 μm/min, indicating that the microtubules are moving poleward. (D) Microtubule length in the metaphase spindle does not change significantly because new tubulin subunits are added at the microtubule plus end at the same rate as tubulin subunits are removed from the minus end. (B and C, from T.J. Mitchison and E.D. Salmon, *Nat. Cell Biol.* 3:E17–21, 2001. With permission from Macmillan Publishers Ltd.)

One of the most striking aspects of mitosis in vertebrate cells is the continuous oscillatory movement of the chromosomes in prometaphase and metaphase. When studied by video microscopy in newt lung cells, the movements are seen to switch between two states—a poleward (P) state, when the chromosomes are pulled toward the pole, and an away-from-the-pole (AP), or neutral, state, when the poleward forces are turned off and the polar ejection force pushes the chromosomes away from the pole. The switch between the two states may depend on the degree of tension in the kinetochore. It has been proposed, for example, that as chromosomes move toward the spindle pole, an increasing polar ejection force generates tension in the kinetochore nearest the

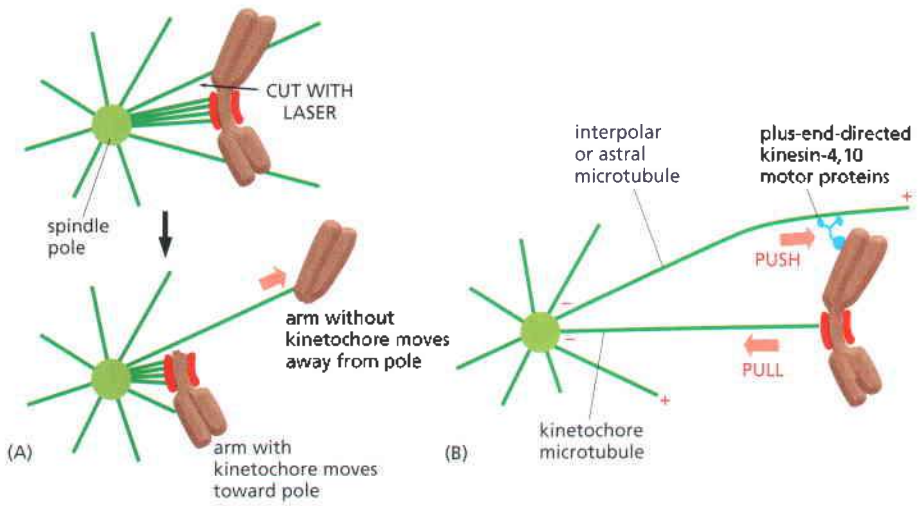


Figure 17–42 How opposing forces may drive chromosomes to the metaphase plate. (A) Evidence for a polar ejection force that pushes chromosomes away from the spindle poles toward the spindle equator. In this experiment, a laser beam severs a prometaphase chromosome that is attached to a single pole by kinetochore microtubules. The part of the severed chromosome without a kinetochore is pushed rapidly away from the pole, whereas the part with the kinetochore moves toward the pole, reflecting a decreased repulsion. (B) A model of how two opposing forces may cooperate to move chromosomes to the metaphase plate. Plus-end-directed motor proteins (kinesin-4 and kinesin-10) on the chromosome arms are thought to interact with microtubules to generate the polar ejection force, which pushes chromosomes toward the spindle equator (see Figure 17–30). Poleward forces generated by depolymerization at the kinetochore, together with microtubule flux, are thought to pull chromosomes toward the pole.

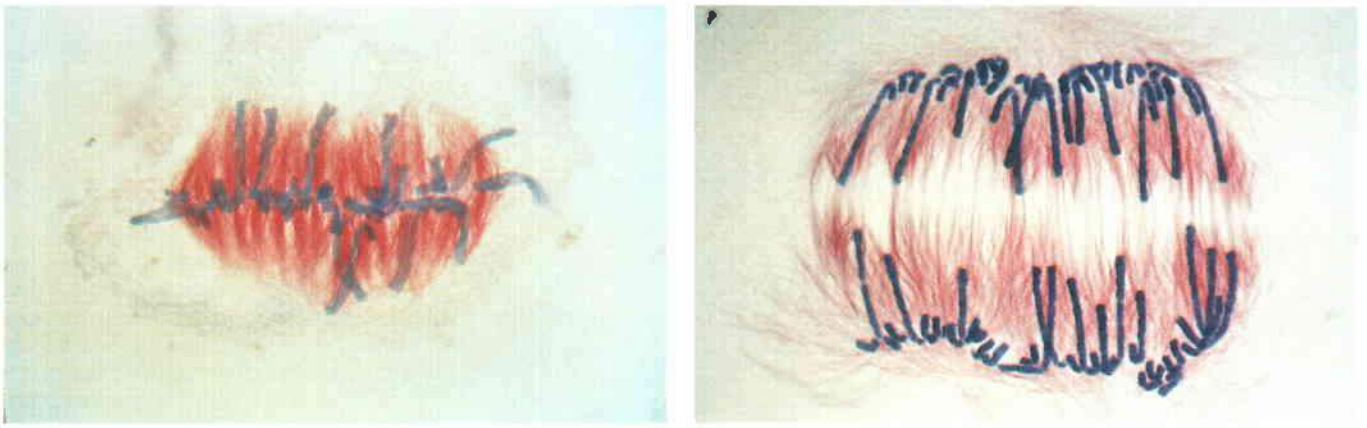


Figure 17–43 Sister-chromatid separation at anaphase. In the transition from metaphase (A) to anaphase (B), sister chromatids suddenly separate and move toward opposite poles of the mitotic spindle—as shown in these light micrographs of *Haemanthus* (lily) endosperm cells that were stained with gold-labeled antibodies against tubulin. (Courtesy of Andrew Bajer.)

pole, triggering a switch to the away-from-the-pole state and gradually resulting in the accumulation of chromosomes at the equator of the spindle.

The APC/C Triggers Sister-Chromatid Separation and the Completion of Mitosis

After M-Cdk has triggered the complex rearrangements that occur in early mitosis, the cell cycle reaches its climax with the separation of the sister chromatids at the metaphase-to-anaphase transition (**Figure 17–43**). Although M-Cdk activity sets the stage for this event, the anaphase-promoting complex (APC/C) discussed earlier throws the switch that initiates sister-chromatid separation by ubiquitylating several mitotic regulatory proteins and thereby triggering their destruction (see Figure 17–20A).

During metaphase, cohesins holding the sister chromatids together resist the poleward forces that pull the sister chromatids apart. Anaphase begins with a sudden disruption of sister-chromatid cohesion, which allows the sisters to separate and move to opposite poles of the spindle. The APC/C initiates the process by targeting the inhibitory protein **securin** for destruction. Before anaphase, securin binds to and inhibits the activity of a protease called **separase**. The destruction of securin at the end of metaphase releases separase, which is then free to cleave one of the subunits of cohesin. The cohesins fall away, and the sister chromatids abruptly and synchronously separate (**Figure 17–44**).

In addition to securin, the APC/C also targets the S- and M-cyclins for destruction, leading to the loss of most Cdk activity in anaphase. Cdk inactivation allows phosphatases to dephosphorylate the many Cdk target substrates in the cell, as required for the completion of mitosis and cytokinesis (discussed later).

If the APC/C triggers anaphase, what activates the APC/C? The answer is only partly known. As mentioned earlier, APC/C activation requires the protein Cdc20, which binds to and activates the APC/C in mitosis (see Figure 17–20A). At least two processes regulate Cdc20 and its association with the APC/C. First, Cdc20 synthesis increases as the cell approaches mitosis, owing to an increase in the transcription of its gene. Second, phosphorylation of the APC/C helps Cdc20 bind to the APC/C, thereby helping to create an active complex. Among the kinases that phosphorylate and thus activate the APC/C is M-Cdk. Thus, M-Cdk not only triggers the early mitotic events leading up to metaphase, but it also sets the stage for progression into anaphase. The ability of M-Cdk to promote Cdc20–APC/C activity creates a negative feedback loop: M-Cdk sets in motion a regulatory process that leads to cyclin destruction and thus its own inactivation.

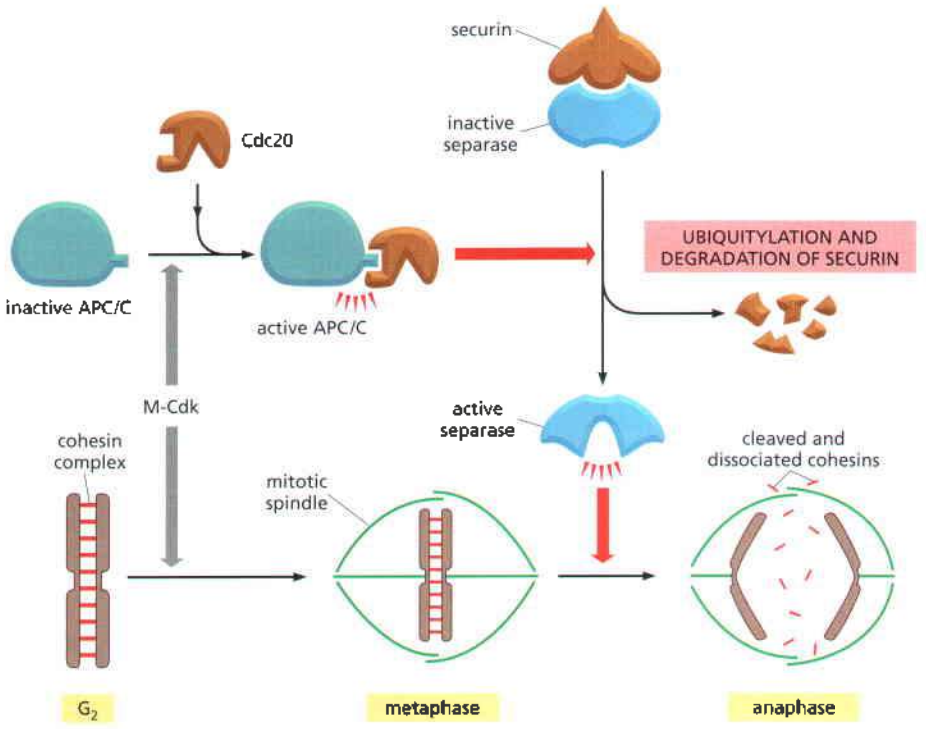


Figure 17–44 The initiation of sister-chromatid separation by the APC/C. The activation of APC/C by Cdc20 leads to the ubiquitylation and destruction of securin, which normally holds separase in an inactive state. The destruction of securin allows separase to cleave Scc1, a subunit of the cohesin complex holding the sister chromatids together (see Figure 17–24). The pulling forces of the mitotic spindle then pull the sister chromatids apart. In animal cells, phosphorylation by Cdks also inhibits separase (not shown). Thus, Cdk inactivation in anaphase (resulting from cyclin destruction) also promotes separase activation by allowing its dephosphorylation.

Unattached Chromosomes Block Sister-Chromatid Separation: The Spindle Assembly Checkpoint

Cells usually spend about half of mitosis in metaphase, with the chromosomes aligned on the metaphase plate, jostling about, awaiting the APC/C signal that induces sister chromatids to separate. Drugs that destabilize microtubules, such as colchicine or vinblastine (discussed in Chapter 16), arrest cells in mitosis for hours or even days. This observation led to the identification of a **spindle assembly checkpoint** mechanism that is activated by the drug treatment and blocks progression through the metaphase-to-anaphase transition. The checkpoint mechanism ensures that cells do not enter anaphase until all chromosomes are correctly bi-oriented on the mitotic spindle.

The spindle assembly checkpoint depends on a sensor mechanism that monitors the strength of microtubule attachment, and possibly tension, at the kinetochore. Any kinetochore that is not properly attached to the spindle sends out a negative signal that blocks Cdc20–APC/C activation and thus blocks the metaphase-to-anaphase transition. Only when the last kinetochore is properly attached is this block removed, allowing sister-chromatid separation to occur.

It is thought that inappropriately attached kinetochores somehow generate a diffusible signal that inhibits Cdc20–APC/C activity throughout the cell. The molecular basis of this signal is not clear, although several proteins, including *Mad2*, are recruited to unattached kinetochores and are required for the spindle assembly checkpoint to function (**Figure 17–45**). One appealing possibility, based primarily on detailed structural analyses of Mad2, is that the unattached kinetochore acts like an enzyme that catalyzes a change in the conformation of Mad2, so that Mad2 can bind and inhibit Cdc20–APC/C.

In mammalian somatic cells, the spindle assembly checkpoint determines the normal timing of anaphase. The destruction of securin in these cells begins moments after the last sister chromatid pair becomes bi-oriented on the spindle, and anaphase begins about 20 minutes later. Experimental inhibition of the checkpoint mechanism causes premature sister-chromatid separation and anaphase. Surprisingly, the normal timing of anaphase does not depend on the spindle assembly checkpoint in some cells, such as yeasts and the cells of early frog and fly embryos. Some other mechanism, as yet unknown, must determine the timing of anaphase in these cells.

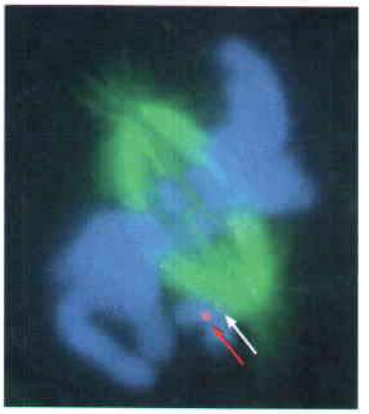
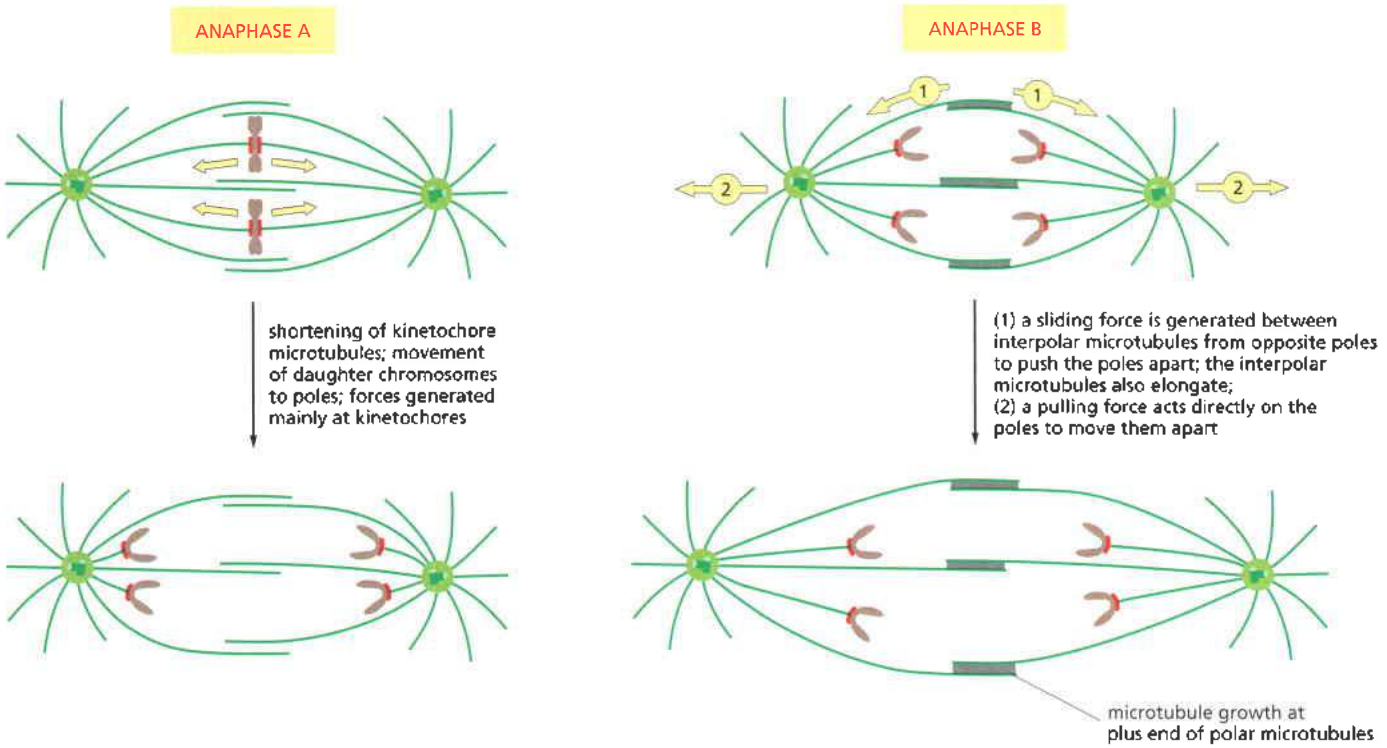


Figure 17–45 Mad2 protein on unattached kinetochores. This fluorescence micrograph shows a mammalian cell in prometaphase, with the mitotic spindle in *green* and the sister chromatids in *blue*. One sister chromatid pair is attached to only one pole of the spindle. Staining with anti-Mad2 antibodies indicates that Mad2 is bound to the kinetochore of the unattached sister chromatid (*red dot*, indicated by *red arrow*). A small amount of Mad2 is associated with the kinetochore of the sister chromatid that is attached to the spindle pole (*pale dot*, indicated by *white arrow*). (From J.C. Waters et al., *J. Cell Biol.* 141:1181–1191, 1998. With permission from The Rockefeller University Press.)



Chromosomes Segregate in Anaphase A and B

The sudden loss of sister-chromatid cohesion at the onset of anaphase leads to sister-chromatid separation, which allows the forces of the mitotic spindle to pull the sisters to opposite poles of the cell—called *chromosome segregation*. The chromosomes move by two independent and overlapping processes. The first, referred to as **anaphase A**, is the initial poleward movement of the chromosomes, which is accompanied by shortening of the kinetochore microtubules. The second, referred to as **anaphase B**, is the separation of the spindle poles themselves, which begins after the sister chromatids have separated and the daughter chromosomes have moved some distance apart (**Figure 17–46**).

Chromosome movement in anaphase A depends on a combination of the two major poleward forces described earlier. The first is the force generated by microtubule depolymerization at the kinetochore, which results in the loss of tubulin subunits at the plus end as the kinetochore moves toward the pole. The second is provided by microtubule flux, which is the poleward movement of the microtubules toward the spindle pole, where minus end depolymerization occurs. The relative importance of these two forces during anaphase varies in different cell types: in embryonic cells, chromosome movement depends mainly on microtubule flux, for example, whereas movement in yeast and vertebrate somatic cells results primarily from forces generated at the kinetochore.

Spindle pole separation during anaphase B depends on motor-driven mechanisms similar to those that separate the two centrosomes in early mitosis (see Figure 17–30). Plus-end directed kinesin-5 motor proteins, which cross-link the overlapping plus ends of the interpolar microtubules, push the poles apart. In addition, dynein motors that anchor astral microtubule plus ends to the cell cortex pull the poles apart.

Although sister-chromatid separation initiates the chromosome movements of anaphase A, other mechanisms also ensure correct chromosome movements in anaphase A and spindle elongation in anaphase B. Most importantly, the completion of a normal anaphase depends on the dephosphorylation of Cdk substrates, which in most cells results from the APC/C-dependent destruction of cyclins. If M-cyclin destruction is prevented—by the production of a mutant form that is not recognized by the APC/C, for example—sister-chromatid separation generally occurs but the chromosome movements and microtubule behavior of anaphase are abnormal.

Figure 17–46 The major forces that separate sister chromatids at anaphase in mammalian cells. Chromosome movement toward the poles in anaphase A depends on the depolymerization of kinetochore microtubules and poleward microtubule flux. In anaphase B, the two spindle poles move apart. Two separate forces are thought to be responsible for anaphase B: the elongation and sliding of the interpolar microtubules past one another in the central spindle push the two poles apart, and motor proteins attached to the plasma membrane near each spindle pole act on astral microtubules to pull the poles away from each other, toward the cell surface.

The relative contributions of anaphase A and anaphase B to chromosome segregation vary greatly, depending on the cell type. In mammalian cells, anaphase B begins shortly after anaphase A and stops when the spindle is about twice its metaphase length; in contrast, the spindles of yeasts and certain protozoa primarily use anaphase B to separate the chromosomes at anaphase, and their spindles elongate to up to 15 times the metaphase length in the process.

Segregated Chromosomes Are Packaged in Daughter Nuclei at Telophase

By the end of anaphase, the daughter chromosomes have segregated into two equal groups at opposite ends of the cell. **In telophase**, the final stage of mitosis, the two sets of chromosomes are packaged into a pair of daughter nuclei. The first major event of telophase is the disassembly of the mitotic spindle, followed by the re-formation of the nuclear envelope. Initially, nuclear membrane fragments associate with the surface of individual chromosomes. These membrane fragments fuse to partly enclose clusters of chromosomes and then coalesce to re-form the complete nuclear envelope. Nuclear pore complexes are incorporated into the envelope, the nuclear lamina re-forms, and the envelope once again becomes continuous with the endoplasmic reticulum. Once the nuclear envelope has re-formed, the pore complexes pump in nuclear proteins, the nucleus expands, and the condensed mitotic chromosomes are reorganized into their interphase state, allowing gene transcription to resume. A new nucleus has been created, and mitosis is complete. All that remains is for the cell to complete its division into two.

We saw earlier that phosphorylation of various proteins by M-Cdk promotes spindle assembly, chromosome condensation, and nuclear envelope breakdown in early mitosis. It is thus not surprising that the dephosphorylation of these same proteins is required for spindle disassembly and the re-formation of daughter nuclei in telophase. In principle, these dephosphorylations and the completion of mitosis could be triggered by the inactivation of Cdks, the activation of phosphatases, or both. Although Cdk inactivation—resulting primarily from cyclin destruction—is mainly responsible in most cells, some cells also rely on activation of phosphatases. In budding yeast, for example, the completion of mitosis depends on the activation of a phosphatase called *Cdc14*, which dephosphorylates a subset of Cdk substrates involved in anaphase and telophase.

Meiosis Is a Special Form of Nuclear Division Involved in Sexual Reproduction

Most eucaryotic organisms reproduce sexually: the genomes of two parents mix to generate offspring that are genetically distinct from either parent (discussed in Chapter 21). The cells of these organisms are generally *diploid*: that is, they contain two slightly different copies, or *homologs*, of each chromosome, one from each parent. Sexual reproduction depends on a specialized nuclear division process called *meiosis*, which produces *haploid* cells carrying only a single copy of each chromosome. In many organisms, the haploid cells differentiate into specialized reproductive cells called *gametes*—eggs and sperm in most species. In these species, the reproductive cycle ends when a sperm and egg fuse to form a diploid zygote with the potential to form a new individual. Here, we consider the basic mechanisms and regulation of meiosis, with an emphasis on how they compare with those of mitosis. Meiosis is discussed in more detail in Chapter 21.

Meiosis begins with a round of chromosome duplication, called *meiotic S phase*, followed by two rounds of chromosome segregation, called *meiosis I* and *II*. Meiosis I segregates the homologs (each composed of a tightly linked pair of sister chromatids). Meiosis II, like conventional mitosis, segregates the sister chromatids of each homolog (**Figure 17–47**).

The first meiotic division solves the central problem of meiosis: how to segregate the homologous chromosomes. Like sister-chromatid segregation in

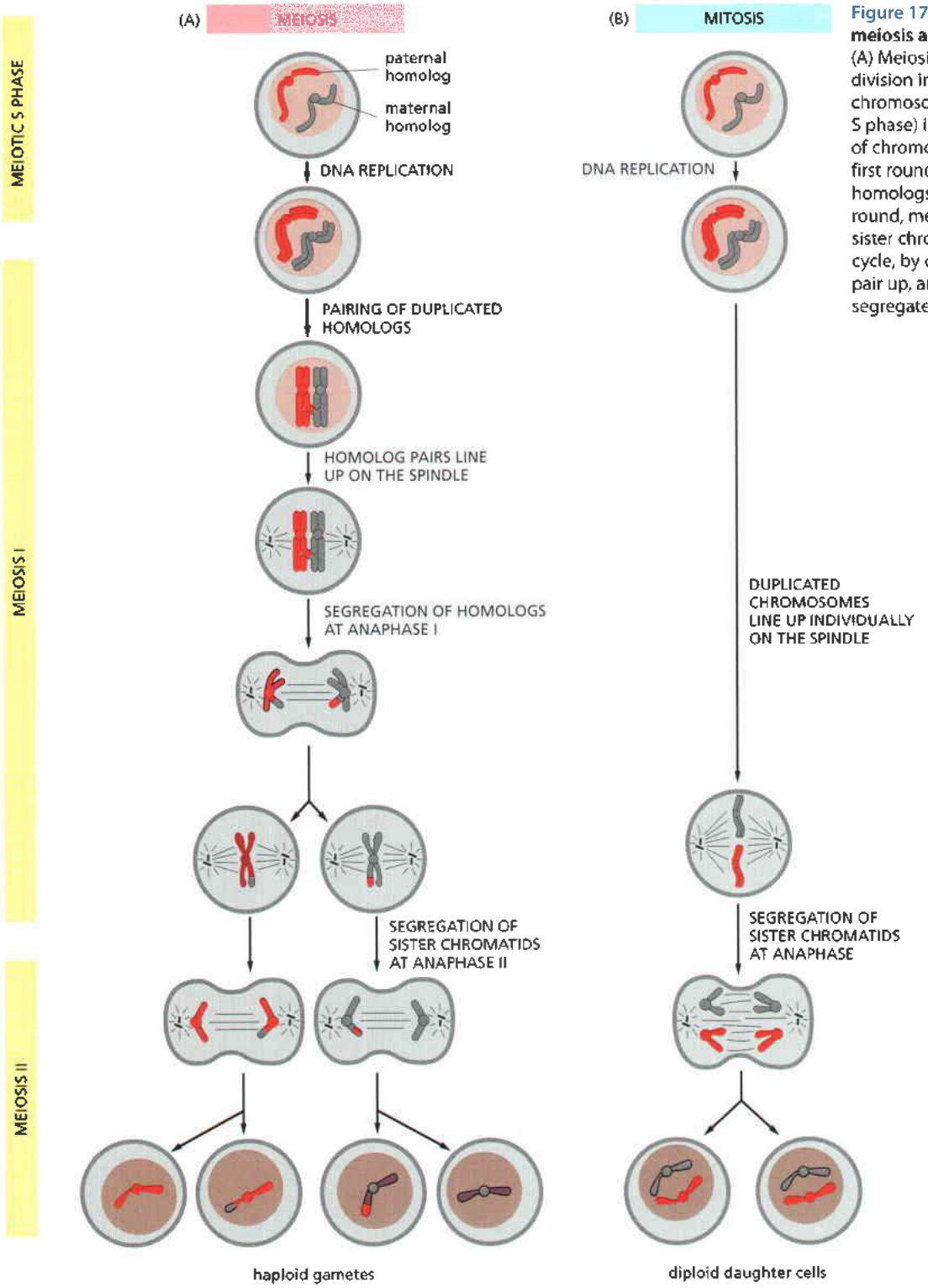


Figure 17–47 Comparison of meiosis and the mitotic cell cycle. (A) Meiosis is a form of nuclear division in which a single round of chromosome duplication (meiotic S phase) is followed by two rounds of chromosome segregation. The first round, meiosis I, segregates the homologs, whereas the second round, meiosis II, segregates the sister chromatids. (B) In a mitotic cell cycle, by contrast, homologs do not pair up, and the sister chromatids segregate in a single division.

mitosis, homolog segregation in meiosis I depends on the formation of linkages between homologs. These linkages allow the homolog pairs to be bi-oriented on the first meiotic spindle, with the homologs in a pair attached to opposite poles. Homolog linkage is removed at the onset of anaphase I, allowing the spindle to pull the homologs to opposite ends of the cell.

Linkages form between homologs by a remarkably complex and lengthy process that occurs after meiotic S phase, in a period called *meiotic prophase* or *prophase I*. This process begins with homolog *pairing*, whereby the homologs

gradually move close to each other in the nucleus, primarily as a result of interactions between complementary DNA sequences in the two homologs. Homolog linkages are then locked in place by homologous recombination between nonsister chromatids in each homolog pair: DNA double-strand breaks are formed at several locations in each sister chromatid, resulting in large numbers of DNA recombination events between the homologs. Some of these events lead to reciprocal DNA exchanges called *crossovers*, where the DNA of a chromatid crosses over to become continuous with the DNA of a homologous chromatid (**Figure 17–48**). At least one of these crossovers occurs in each homolog pair, ensuring that the homologs in every pair are physically connected when the cell enters the first meiotic division.

Another uniquely meiotic problem must be solved when the homolog pairs are attached to the first meiotic spindle. Each homolog contains two tightly linked sister chromatids, and thus the attachment of a homolog to a spindle pole requires that both sister kinetochores in a homolog attach to the same pole. This type of attachment is normally avoided during mitosis (see Figure 17–39). In meiosis I, however, the two sister kinetochores are somehow fused into a single microtubule-binding unit that attaches to just one pole (see Figure 21–12A). These mechanisms are reversed after meiosis I, so that in meiosis II the sister chromatid pairs can be bi-oriented on the spindle as they are in mitosis.

Crossovers hold homolog pairs together only because the arms of the sister chromatids are connected by sister-chromatid cohesion (see Figure 17–48). The loss of cohesion from sister-chromatid arms therefore triggers homolog separation at the onset of anaphase I. In most species, the loss of arm cohesion in meiosis I depends on APC/C activation, which leads to securin destruction, separase activation, and cohesin cleavage along the arms (see Figure 17–44). In contrast to mitosis, however, cohesin complexes near the centromeres remain uncleaved in meiosis I because cohesin in that region is protected from separase (discussed in Chapter 21). Sister-chromatid pairs therefore remain linked at their centromeres throughout meiosis I, allowing their correct bi-orientation on the spindle in meiosis II. The mechanisms that block cohesin cleavage at the centromere in meiosis I are removed in meiosis II. At the onset of anaphase II, APC/C activation therefore triggers centromeric cohesin cleavage and sister-chromatid separation—much as it does in mitosis.

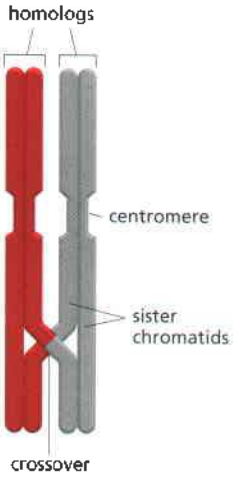


Figure 17–48 A crossover between homologs. As in mitosis, the sister chromatids in each homolog are tightly connected along their entire lengths. A single crossover has occurred between two nonsister chromatids in this example, but either of the two chromatids of a homolog can form a crossover with either chromatid of the other homolog, and it is usual for multiple crossovers to be formed.

Summary

M-Cdk triggers the events of early mitosis, including chromosome condensation, assembly of the mitotic spindle, and bipolar attachment of the sister chromatid pairs to microtubules of the spindle. Spindle formation in animal cells depends largely on the ability of mitotic chromosomes to stimulate local microtubule nucleation and stability, as well as on the ability of motor proteins to organize microtubules into a bipolar array. Many cells also use centrosomes to facilitate spindle assembly. Anaphase is triggered by the APC/C, which stimulates the destruction of the proteins that hold the sister chromatids together. APC/C also promotes cyclin destruction and thus the inactivation of M-Cdk. The resulting dephosphorylation of Cdk targets is required for the events that complete mitosis, including the disassembly of the spindle and the re-formation of the nuclear envelope. Meiosis is a specialized form of nuclear division in which a single round of chromosome duplication is followed by two rounds of chromosome segregation, resulting in the formation of haploid nuclei.

CYTOKINESIS

The final step in the cell cycle is **cytokinesis**, the division of the cytoplasm. In a typical cell, cytokinesis accompanies every mitosis, although some cells, such as early *Drosophila* embryos (discussed later) and some mammalian hepatocytes and heart muscle cells, undergo mitosis without cytokinesis and thereby acquire multiple nuclei. In most animal cells, cytokinesis begins in anaphase and ends shortly after the completion of mitosis in telophase.

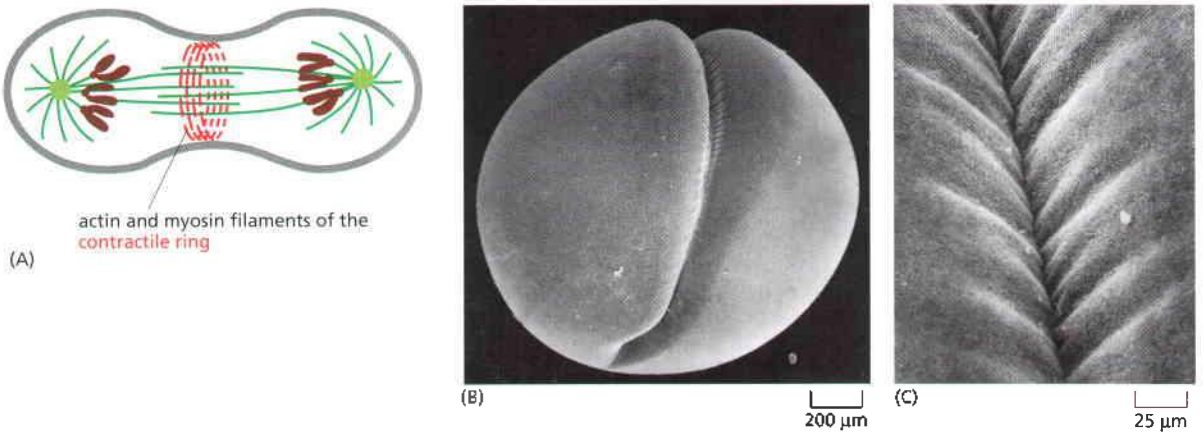


Figure 17–49 **Cytokinesis.** (A) The actin-myosin bundles of the contractile ring are oriented as shown, so that their contraction pulls the membrane inward. (B) In this low-magnification scanning electron micrograph of a cleaving frog egg, the cleavage furrow is especially obvious and well defined, as the cell is unusually large. The furrowing of the cell membrane is caused by the activity of the contractile ring underneath it. (C) The surface of a furrow at higher magnification. (B and C, from H.W. Beams and R.G. Kessel, *Am. Sci.* 64:279–290, 1976. With permission from Sigma Xi.)

The first visible change of cytokinesis in an animal cell is the sudden appearance of a pucker, or *cleavage furrow*, on the cell surface. The furrow rapidly deepens and spreads around the cell until it completely divides the cell in two. In animal cells and many unicellular eucaryotes, the structure underlying this process is the *contractile ring*—a dynamic assembly composed of actin filaments, myosin II filaments, and many structural and regulatory proteins. During anaphase, the ring assembles just beneath the plasma membrane (**Figure 17–49**; see also Panel 17–1). The ring gradually contracts, and, at the same time, fusion of intracellular vesicles with the plasma membrane inserts new membrane adjacent to the ring. This addition of membrane compensates for the increase in surface area that accompanies cytoplasmic division. When ring contraction is completed, membrane insertion and fusion seal the gap between the daughter cells. Thus, cytokinesis can be considered to occur in four stages—initiation, contraction, membrane insertion, and completion.

Actin and Myosin II in the Contractile Ring Generate the Force for Cytokinesis

In interphase cells, actin and myosin filaments form a cortical network underlying the plasma membrane. In some cells, they also form large cytoplasmic bundles called *stress fibers* (discussed in Chapter 16). As cells enter mitosis, these arrays of actin and myosin disassemble; much of the actin reorganizes, and myosin II filaments are released. As the sister chromatids separate in anaphase, actin and myosin II begin to accumulate in the rapidly assembling **contractile ring** (**Figure 17–50**), which also contains numerous other proteins that provide

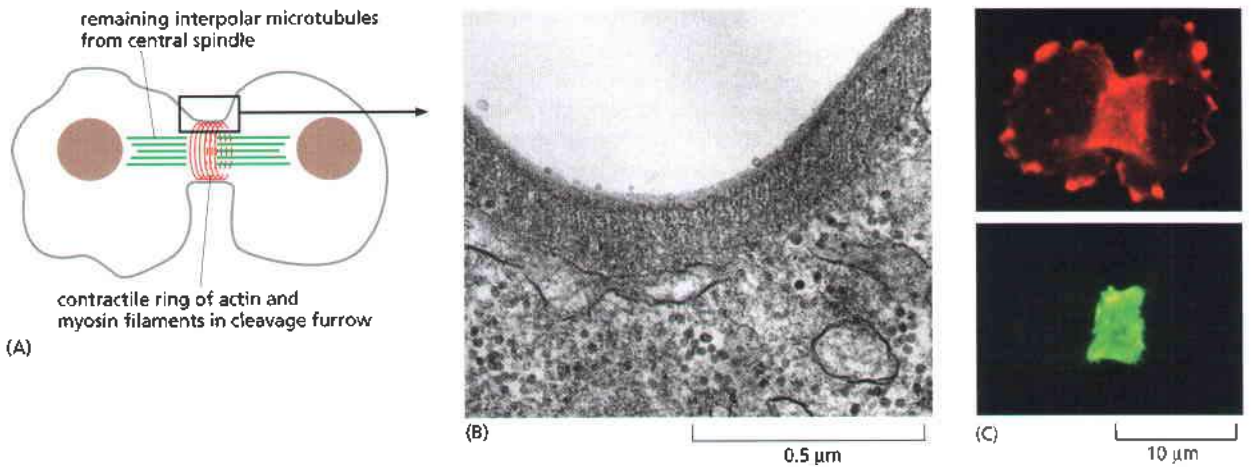


Figure 17–50 **The contractile ring.** (A) A drawing of the cleavage furrow in a dividing cell. (B) An electron micrograph of the ingrowing edge of a cleavage furrow of a dividing animal cell. (C) Fluorescence micrographs of a dividing slime mold amoeba stained for actin *(red)* and myosin II *(green)*. Whereas all of the visible myosin II has redistributed to the contractile ring, only some of the actin has done so; the rest remains in the cortex of the nascent daughter cells. (B, from H.W. Beams and R.G. Kessel, *Am. Sci.* 64:279–290, 1976. With permission from Sigma Xi; C, courtesy of Yoshio Fukui.)

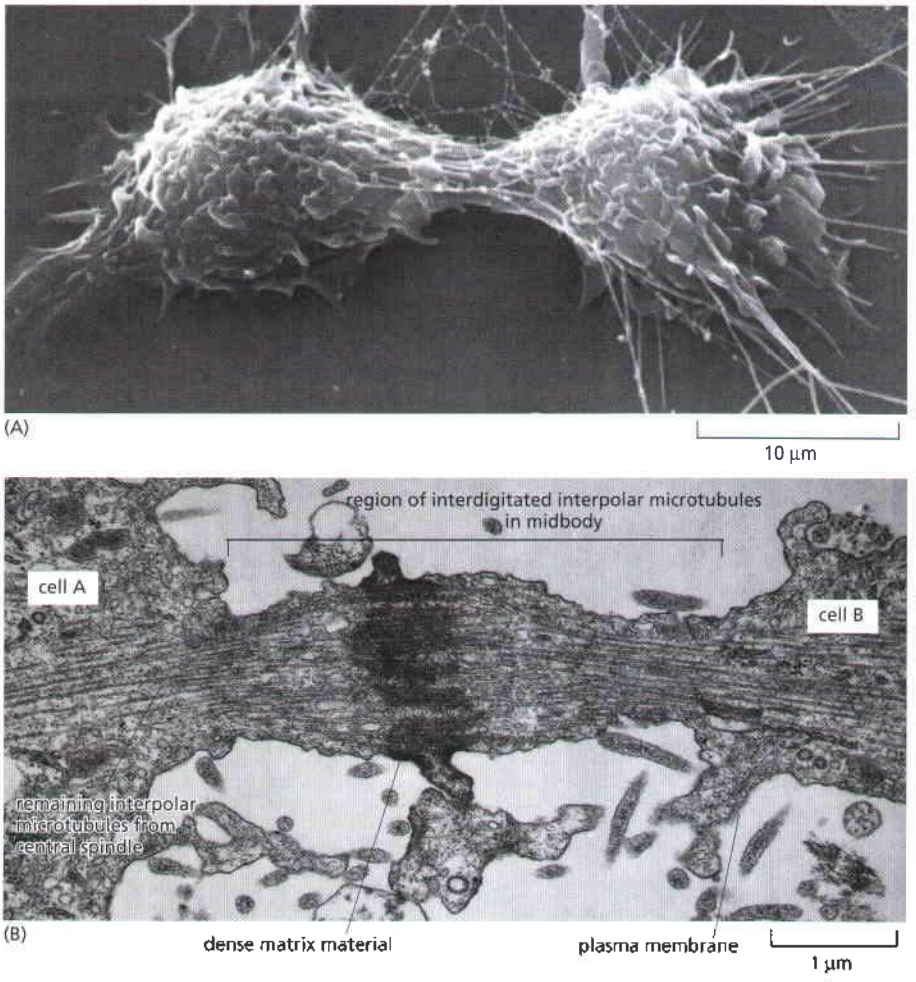


Figure 17–51 **The midbody.** (A) A scanning electron micrograph of a cultured animal cell in the process of dividing; the midbody still joins the two daughter cells. (B) A conventional electron micrograph of the midbody of a dividing animal cell. Cleavage is almost complete, but the daughter cells remain attached by this thin strand of cytoplasm containing the remains of the central spindle. (A, courtesy of Guenter Albrecht-Buehler; B, courtesy of J.M. Mullins.)

structural support or assist in ring assembly. Assembly of the contractile ring results in part from the local formation of new actin filaments, which depends on *formin* proteins that nucleate the assembly of parallel arrays of linear, unbranched actin filaments (discussed in Chapter 16). After anaphase, the overlapping arrays of actin and myosin II filaments contract to generate the force that divides the cytoplasm in two. Once contraction begins, the ring exerts a force large enough to bend a fine glass needle that is inserted in its path. As the ring constricts, it maintains the same thickness, suggesting that its total volume and the number of filaments it contains decrease steadily. Moreover, unlike actin in muscle, the actin filaments in the ring are highly dynamic, and their arrangement changes continually during cytokinesis.

The contractile ring is finally dispensed with altogether when cleavage ends, as the plasma membrane of the cleavage furrow narrows to form the **midbody**. The midbody persists as a tether between the two daughter cells and contains the remains of the central spindle, a large protein structure derived from the antiparallel interpolar microtubules of the spindle midzone, packed tightly together within a dense matrix material (**Figure 17–51**). After the daughter cells separate completely, some of the components of the residual midbody often remain on the inside of the plasma membrane of each cell, where they may serve as a mark on the cortex that helps to orient the spindle in the subsequent cell division.

Local Activation of RhoA Triggers Assembly and Contraction of the Contractile Ring

RhoA, a small GTPase of the Ras superfamily (see Table 15–5), controls the assembly and function of the contractile ring at the site of cleavage. RhoA is activated at

the cell cortex at the future division site, where it promotes actin filament formation, myosin II assembly, and ring contraction. It promotes actin filament formation by activating formins, and it promotes myosin II assembly and contractions by activating multiple protein kinases, including the Rho-activated kinase Rock (Figure 17–52). These kinases phosphorylate the regulatory myosin light chain (RMLC), which is one of the subunits of myosin II. Phosphorylation of the RMLC stimulates bipolar myosin II filament formation and motor activity, thereby promoting the assembly and contraction of the actin–myosin ring.

Like other GTPases, RhoA is inactive when bound to GDP and active when bound to GTP (discussed in Chapter 15). The local activation of RhoA at the cleavage furrow is thought to depend on a Rho guanine nucleotide exchange factor (RhoGEF), which is found at the cell cortex at the future division site and stimulates the release of GDP and binding of GTP to RhoA. We know little about how the RhoGEF is localized or activated at the division site, although the microtubules of the anaphase spindle seem to be involved, as we discuss next.

The Microtubules of the Mitotic Spindle Determine the Plane of Animal Cell Division

The central problem in cytokinesis is how to ensure that division occurs at the right time and in the right place. Cytokinesis must occur only after the two sets of chromosomes are fully segregated from each other, and the site of division must be placed between the two sets of daughter chromosomes, thereby ensuring that each daughter cell receives a complete set. The correct timing and positioning of cytokinesis in animal cells are achieved by elegant mechanisms that depend on the mitotic spindle. During anaphase, the spindle generates signals that initiate furrow formation at a position midway between the spindle poles, thereby ensuring that division occurs between the two sets of separated chromosomes. Because these signals originate in the anaphase spindle, this mechanism also contributes to the correct timing of cytokinesis in late mitosis. Cytokinesis also occurs at the correct time because dephosphorylation of Cdk

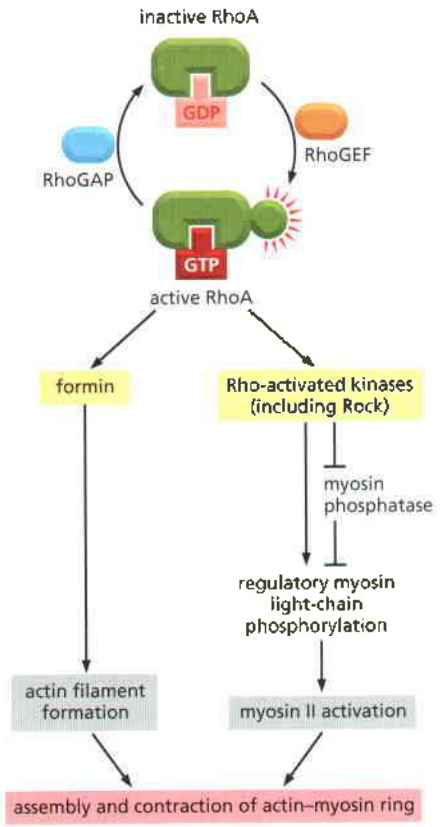


Figure 17–52 **Regulation of the contractile ring by the GTPase RhoA.** Like other Rho family GTPases, RhoA is activated by a RhoGEF protein and inactivated by a RhoGAP protein. The active GTP-bound form of RhoA is focused at the future cleavage site. By binding formins, activated RhoA promotes the assembly of actin filaments in the contractile ring. By activating Rho-activated protein kinases, such as Rock, it stimulates myosin II filament formation and activity, thereby promoting contraction of the ring.

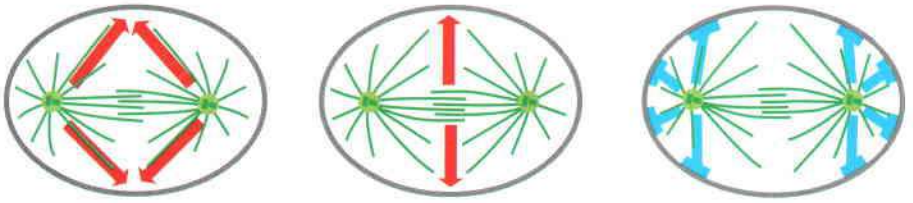


Figure 17–53 Three current models of how the microtubules of the anaphase spindle generate signals that influence the positioning of the contractile ring. No single model explains all the observations, and it is likely that furrow positioning is determined by a combination of these mechanisms, with the importance of the different mechanisms varying in different organisms.

substrates, which depends on cyclin destruction in metaphase and anaphase, initiates cytokinesis. We now describe these regulatory mechanisms in more detail, with an emphasis on cytokinesis in animal cells.

Studies of the fertilized eggs of marine invertebrates first revealed the importance of spindle microtubules in determining the placement of the contractile ring. After fertilization, these embryos cleave rapidly without intervening periods of growth. In this way, the original egg is progressively divided up into smaller and smaller cells. Because the cytoplasm is clear, the spindle can be observed in real time with a microscope. If the spindle is tugged into a new position with a fine glass needle in early anaphase, the incipient cleavage furrow disappears, and a new one develops in accord with the new spindle site—supporting the idea that signals generated by the spindle induce local furrow formation.

How does the mitotic spindle specify the site of division? Three general mechanisms have been proposed, and most cells appear to employ a combination of these (**Figure 17–53**). The first is termed the *astral stimulation model*, which postulates that the astral microtubules carry furrow-inducing signals to the cell cortex, where they are somehow focused into a ring halfway between the spindle poles. Evidence for this model comes from ingenious experiments in large embryonic cells, which demonstrate that a cleavage furrow forms midway between two asters, even when the two centrosomes nucleating the asters are not connected to each other by a mitotic spindle (**Figure 17–54**).

A second possibility, called the *central spindle stimulation model*, is that the spindle midzone, or central spindle, generates a furrow-inducing signal that specifies the site of furrow formation at the cell cortex. The overlapping interpolar microtubules of the central spindle associate with numerous signaling proteins, including proteins that may stimulate RhoA (**Figure 17–55**). Defects in the function of these proteins (in *Drosophila* mutants, for example) result in failure of cytokinesis.

A third model proposes that, in some cell types, the astral microtubules promote the local relaxation of actin–myosin bundles at the cell cortex. According to this *astral relaxation model*, the cortical relaxation is minimal at the spindle equator, thus promoting cortical contraction at that site. In the early embryos of

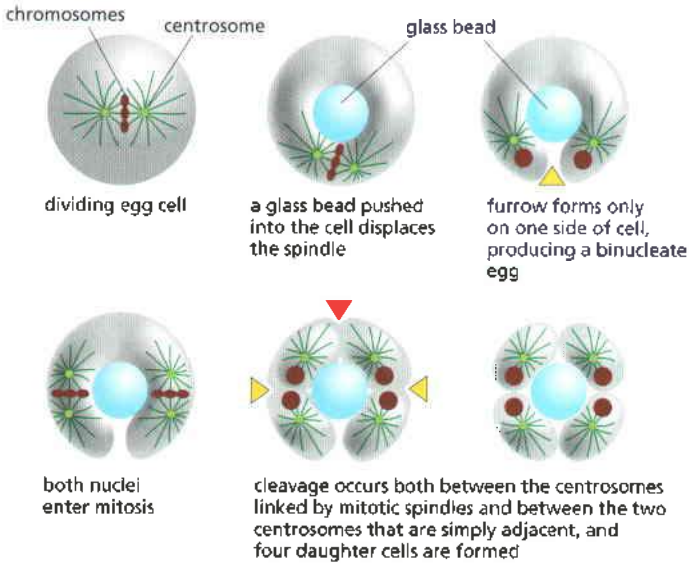


Figure 17–54 An experiment demonstrating the influence of the position of microtubule asters on the subsequent plane of cleavage in a large egg cell. If the mitotic spindle is mechanically pushed to one side of the cell with a glass bead, the membrane furrowing is incomplete, failing to occur on the opposite side of the cell. Subsequent cleavages occur not only at the midzone of each of the two subsequent mitotic spindles *(yellow arrowheads)*, but also between the two adjacent asters that are not linked by a mitotic spindle—but in this abnormal cell share the same cytoplasm *(red arrowhead)*. Apparently, the contractile ring that produces the cleavage furrow in these cells always forms in the region midway between two asters, suggesting that the asters somehow alter the adjacent region of cell cortex to induce furrow formation between them.

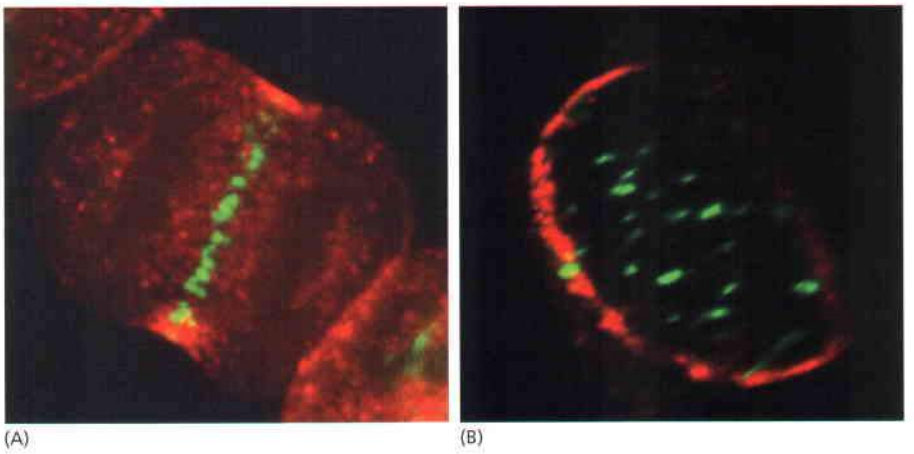


Figure 17–55 **Localization of cytokinesis regulators at the central spindle of the human cell.** (A) Fluorescence micrograph of a cultured human cell at the beginning of cytokinesis reveals the locations of the GTPase RhoA *(red)* and a protein called Cyk4 *(green)*, which is one of several regulatory proteins that form complexes at the overlapping plus ends of interpolar microtubules. These proteins are thought to generate signals that help control RhoA activity at the cell cortex (see Figures 17–52 and 17–53B). (B) When the cell is cross-sectioned in the plane of the contractile ring as shown here, RhoA *(red)* forms a ring beneath the cell surface, while the central spindle protein Cyk4 *(green)* is associated with microtubule bundles scattered throughout the equatorial plane of the cell. (Courtesy of Alisa Piekny and Michael Glotzer.)

C. elegans, for example, treatments that result in the loss of astral microtubules lead to increased contractile activity throughout the cell cortex, consistent with this model.

In some cell types, the site of ring assembly is chosen before mitosis. In budding yeasts, for example, a ring of proteins called *septins* assembles in late G_1 at the future division site. The septins are thought to form a scaffold onto which other components of the contractile ring, including myosin II, assemble. In plant cells, an organized band of microtubules and actin filaments, called the **preprophase band**, assembles just before mitosis and marks the site where the cell wall will assemble and divide the cell in two, as we now discuss.

The Phragmoplast Guides Cytokinesis in Higher Plants

In most animal cells, the inward movement of the cleavage furrow depends on an increase in the surface area of the plasma membrane. New membrane is added at the inner edge of the cleavage furrow and is generally provided by small membrane vesicles that are transported on microtubules from the Golgi apparatus to the furrow.

Membrane deposition is particularly important for cytokinesis in higher-plant cells. These cells are enclosed by a semirigid *cell wall*. Rather than a contractile ring dividing the cytoplasm from the outside in, the cytoplasm of the plant cell is partitioned from the inside out by the construction of a new cell wall, called the **cell plate**, between the two daughter nuclei (**Figure 17–56**). The assembly of the cell plate begins in late anaphase and is guided by a structure called the **phragmoplast**, which contains microtubules derived from the mitotic

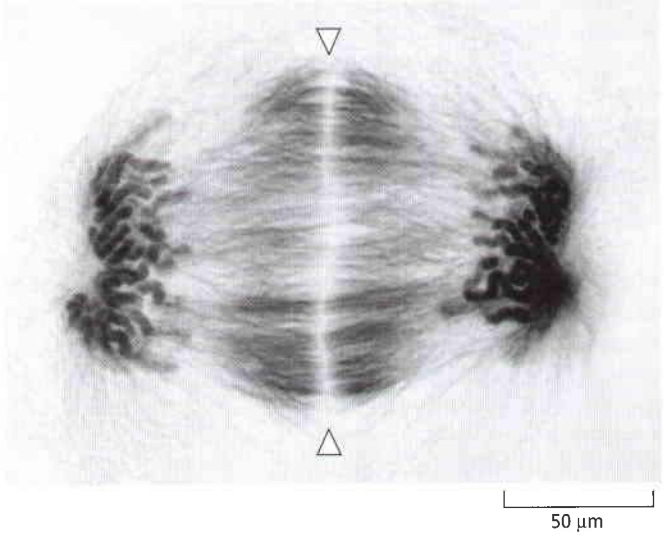


Figure 17–56 **Cytokinesis in a plant cell in telophase.** In this light micrograph, the early cell plate (between the two *arrowheads)* has formed in a plane perpendicular to the plane of the page. The microtubules of the spindle are stained with gold-labeled antibodies against tubulin, and the DNA in the two sets of daughter chromosomes is stained with a fluorescent dye. Note that there are no astral microtubules, because there are no centrosomes in higher-plant cells. (Courtesy of Andrew Bajer.)

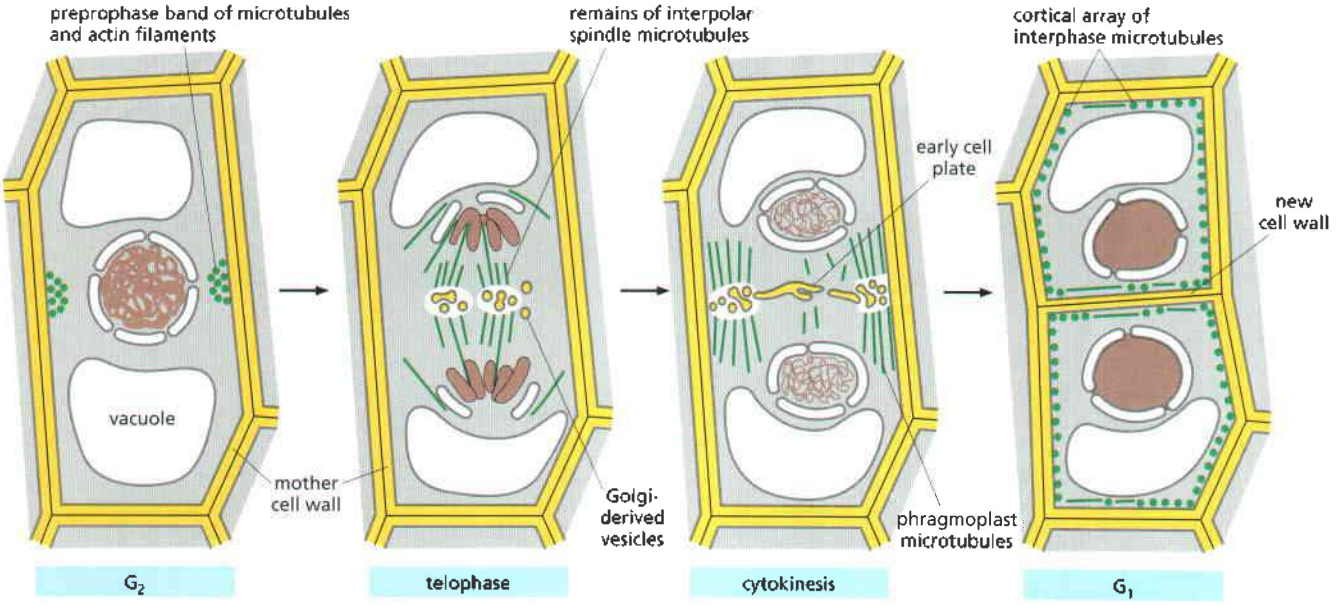


Figure 17–57 The special features of cytokinesis in a higher plant cell. The division plane is established before M phase by a band of microtubules and actin filaments (the preprophase band) at the cell cortex. At the beginning of telophase, after the chromosomes have segregated, a new cell wall starts to assemble inside the cell at the equator of the old spindle. The interpolar microtubules of the mitotic spindle remaining at telophase form the phragmoplast. The plus ends of these microtubules no longer overlap but end at the cell equator. Golgi-derived vesicles, filled with cell-wall material, are transported along these microtubules and fuse to form the new cell wall, which grows outward to reach the plasma membrane and original cell wall. The plasma membrane and the membrane surrounding the new cell wall fuse, completely separating the two daughter cells.

spindle. Motor proteins transport small vesicles along these microtubules from the Golgi apparatus to the cell center. These vesicles, filled with polysaccharide and glycoproteins required for the synthesis of the new cell wall, fuse to form a disclike, membrane-enclosed structure called the *early cell plate*. The plate expands outward by further vesicle fusion until it reaches the plasma membrane and the original cell wall and divides the cell in two. Later, cellulose microfibrils are laid down within the matrix of the cell plate to complete the construction of the new cell wall (**Figure 17–57**).

Membrane-Enclosed Organelles Must Be Distributed to Daughter Cells During Cytokinesis

The process of mitosis ensures that each daughter cell receives a full complement of chromosomes. When a eucaryotic cell divides, however, each daughter cell must also inherit all of the other essential cell components, including the membrane-enclosed organelles. As discussed in Chapter 12, organelles such as mitochondria and chloroplasts cannot be assembled de novo from their individual components; they can arise only by the growth and division of the preexisting organelles. Similarly, cells cannot make a new endoplasmic reticulum (ER) unless some part of it is already present.

How, then, do the various membrane-enclosed organelles segregate when a cell divides? Organelles such as mitochondria and chloroplasts are usually present in large enough numbers to be safely inherited if, on average, their numbers roughly double once each cycle. The ER in interphase cells is continuous with the nuclear membrane and is organized by the microtubule cytoskeleton. Upon entry into M phase, the reorganization of the microtubules and breakdown of the nuclear envelope releases the ER. In most cells, the ER remains largely intact and is cut in two during cytokinesis. The Golgi apparatus is reorganized and fragmented during mitosis. Golgi fragments associate with the spindle poles and are thereby distributed to opposite ends of the spindle, ensuring that each daughter cell inherits the materials needed to reconstruct the Golgi in telophase.

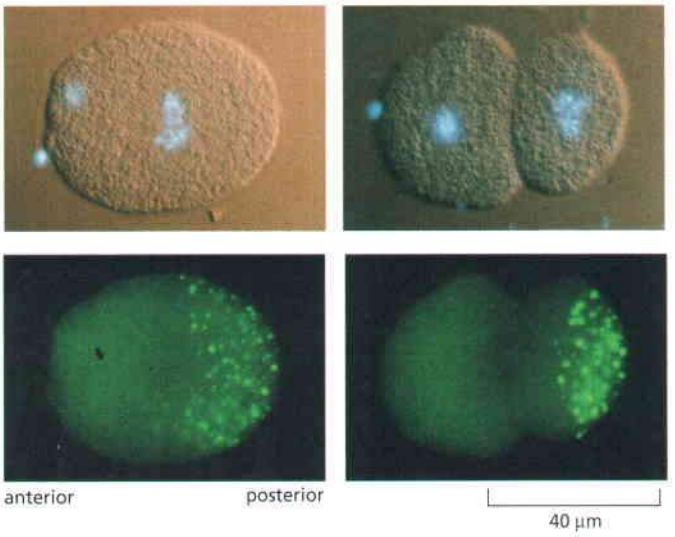


Figure 17–58 An asymmetric cell division segregating cytoplasmic components to only one daughter cell. These light micrographs illustrate the controlled asymmetric segregation of specific cytoplasmic components to one daughter cell during the first division of a fertilized egg of the nematode *C. elegans*. The cells *above* have been stained with a *blue*, DNA-binding, fluorescent dye to show the nucleus (and polar bodies); they are viewed by both differential-interference-contrast and fluorescence microscopy. The cells *below* are the same cells stained with an antibody against P-granules and viewed by fluorescence microscopy. These small granules are made of RNA and proteins and determine which cells become the germ cells. They are distributed randomly throughout the cytoplasm of the unfertilized egg (not shown) but become segregated to the posterior pole of the fertilized egg, as shown on the *left*. The cleavage plane is oriented to ensure that only the posterior daughter cell receives the P-granules when the egg divides, as shown on the *right*. The same segregation process is repeated in several subsequent cell divisions, so that the P-granules end up only in the cells that give rise to eggs and sperm. (Courtesy of Susan Strome.)

Some Cells Reposition Their Spindle to Divide Asymmetrically

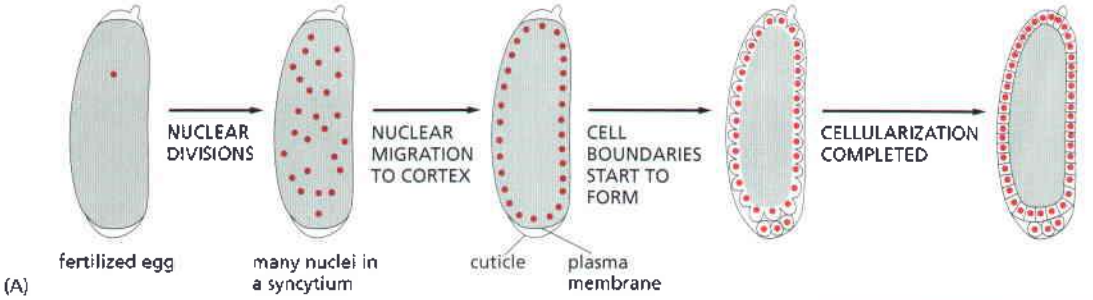
Most animal cells divide symmetrically: the contractile ring forms around the equator of the parent cell, producing two daughter cells of equal size and with the same components. This symmetry results from the placement of the mitotic spindle, which in most cases tends to center itself in the cytoplasm. Astral microtubules and motor proteins that either push or pull on these microtubules contribute to the centering process.

There are many instances in development, however, when cells divide asymmetrically to produce two cells that differ in size, in the cytoplasmic contents they inherit, or in both. Usually, the two different daughter cells are destined to develop along different pathways. To create daughter cells with different fates in this way, the mother cell must first segregate certain components (called *cell fate determinants*) to one side of the cell and then position the plane of division so that the appropriate daughter cell inherits these components (**Figure 17–58**). To position the plane of division asymmetrically, the spindle has to be moved in a controlled manner within the dividing cell. It seems likely that changes in local regions of the cell cortex direct such spindle movements and that motor proteins localized there pull one of the spindle poles, via its astral microtubules, to the appropriate region. Genetic analyses in *C. elegans* and *Drosophila* have identified some of the proteins required for such asymmetric divisions (discussed in Chapter 22), and some of these proteins seem to have a similar role in vertebrates.

Mitosis Can Occur Without Cytokinesis

Although nuclear division is usually followed by cytoplasmic division, there are exceptions. Some cells undergo multiple rounds of nuclear division without intervening cytoplasmic division. In the early *Drosophila* embryo, for example, the first 13 rounds of nuclear division occur without cytoplasmic division, resulting in the formation of a single large cell containing several thousand nuclei, arranged in a monolayer near the surface. A cell in which multiple nuclei share the same cytoplasm is called a **syncytium**. This arrangement greatly speeds up early development, as the cells do not have to take the time to go through all the steps of cytokinesis for each division. After these rapid nuclear divisions, membranes are created around each nucleus in one round of coordinated cytokinesis called *cellularization*. The plasma membrane extends inward and, with the help of an actin–myosin ring, pinches off to enclose each nucleus (**Figure 17–59**). <TTCT>

Nuclear division without cytokinesis also occurs in some types of mammalian cells. Megakaryocytes, which produce blood platelets, and some hepatocytes and heart muscle cells, for example, become multinucleated in this way.



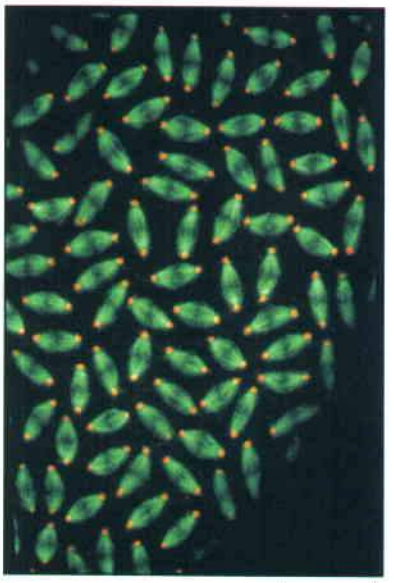


Figure 17–59 Mitosis without cytokinesis in the early *Drosophila* embryo. (A) The first 13 nuclear divisions occur synchronously and without cytoplasmic division to create a large syncytium. Most of the nuclei then migrate to the cortex, and the plasma membrane extends inward and pinches off to surround each nucleus to form individual cells in a process called cellularization. (B) Fluorescence micrograph of multiple mitotic spindles in a *Drosophila* embryo before cellularization. The microtubules are stained *green* and the centrosomes *red*. Note that all the nuclei go through the cycle synchronously; here, they are all in metaphase. (B, courtesy of Kristina Yu and William Sullivan.)

After cytokinesis, most cells enter G_1, in which Cdks are mostly inactive. We end this section by discussing how this state is achieved at the end of M phase.

The G_1 Phase Is a Stable State of Cdk Inactivity

A key regulatory event in late M phase is the inactivation of Cdks, which is driven primarily by APC/C-dependent cyclin destruction. As described earlier in this chapter, the inactivation of Cdks in late M phase has many functions: it triggers the events of late mitosis, promotes cytokinesis, and enables the synthesis of prereplicative complexes at DNA replication origins. It also provides a mechanism for resetting the cell-cycle control system to a state of Cdk inactivity as the cell prepares to enter a new cell cycle. In most cells, this state of Cdk inactivity generates a G_1 gap phase, during which the cell grows and monitors its environment before committing to a new division.

In early animal embryos, the inactivation of M-Cdk in late mitosis is due almost entirely to the action of Cdc20–APC/C, discussed earlier. Recall, however, that M-Cdk stimulates Cdc20–APC/C activity. Thus, the destruction of M-cyclin in late mitosis soon leads to the inactivation of all APC/C activity in an embryonic cell. This APC/C inactivation immediately after mitosis is especially useful in rapid embryonic cell cycles, as it allows the cell to quickly begin accumulating new M-cyclin for the next cycle (**Figure 17–60**A).

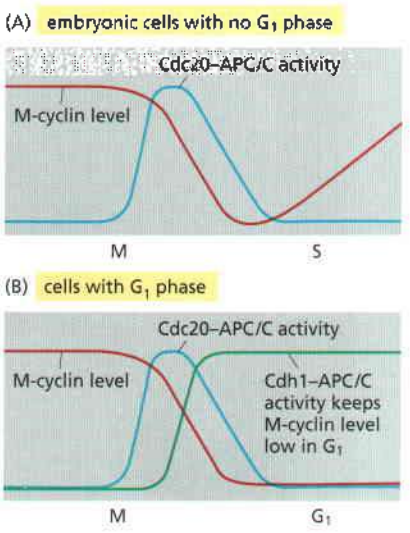


Figure 17–60 The creation of a G_1 phase by stable Cdk inhibition after mitosis. (A) In early embryonic cell cycles, Cdc20–APC/C activity rises at the end of metaphase, triggering M-cyclin destruction. Because M-Cdk activity stimulates Cdc20–APC/C activity, the loss of M-cyclin leads to APC/C inactivation after mitosis, which allows M-cyclins to begin accumulating again. (B) In cells containing a G_1 phase, the drop in M-Cdk activity in late mitosis leads to the activation of Cdh1–APC/C (as well as to the accumulation of Cdk inhibitor proteins; not shown). This ensures a continued suppression of Cdk activity after mitosis, as required for a G_1 phase.

Rapid cyclin accumulation immediately after mitosis is not useful, however, for cells with cell cycles containing a G_1 phase. These cells employ several mechanisms to prevent Cdk reactivation after mitosis. One mechanism uses another APC/C-activating protein called Cdh1, a close relative of Cdc20. Although both Cdh1 and Cdc20 bind to and activate the APC/C, they differ in one important respect. Whereas M-Cdk activates the Cdc20–APC/C complex, it inhibits the Cdh1–APC/C complex by directly phosphorylating Cdh1. As a result of this relationship, Cdh1–APC/C activity increases in late mitosis after the Cdc20–APC/C complex has initiated the destruction of M-cyclin. M-cyclin destruction therefore continues after mitosis: although Cdc20–APC/C activity has declined, Cdh1–APC/C activity is high (Figure 17–60B).

A second mechanism that suppresses Cdk activity in G_1 depends on the increased production of CKIs, the Cdk inhibitory proteins discussed earlier. Budding yeast cells, in which this mechanism is best understood, contain a CKI protein called *Sic1*, which binds to and inactivates M-Cdk in late mitosis and G_1. Like Cdh1, Sic1 is inhibited by M-Cdk, which phosphorylates Sic1 during mitosis and thereby promotes its ubiquitylation by SCF. Thus, Sic1 and M-Cdk, like Cdh1 and M-Cdk, inhibit each other. As a result, the decline in M-Cdk activity that occurs in late mitosis causes the Sic1 protein to accumulate, and this CKI helps keep M-Cdk activity low after mitosis. A CKI protein called *p27* (see Figure 17–19/17–19) may serve similar functions in animal cells.

In most cells, decreased transcription of *M-cyclin* genes also inactivates M-Cdks in late mitosis. In budding yeast, for example, M-Cdk promotes the expression of these genes, resulting in a positive feedback loop. This loop is turned off as cells exit from mitosis: the inactivation of M-Cdk by Cdh1 and Sic1 leads to decreased M-cyclin gene transcription and thus decreased M-cyclin synthesis. Gene regulatory proteins that promote the expression of G_1/S- and S-cyclins are also inhibited during G_1.

Thus, Cdh1-APC/C activation, CKI accumulation, and decreased cyclin gene expression act together to ensure that the early G_1 phase is a time when essentially all Cdk activity is suppressed. As in many other aspects of cell-cycle control, the use of multiple regulatory mechanisms makes the suppression system robust, so that it still operates with reasonable efficiency even if one mechanism fails. So how does the cell escape from this stable G_1 state to initiate a new cell cycle? The answer is that G_1/S-Cdk activity, which rises in late G_1, releases all the braking mechanisms that suppress Cdk activity, as we describe in the next section.

Summary

After mitosis completes the formation of a pair of daughter nuclei, cytokinesis finishes the cell cycle by dividing the cell itself. Cytokinesis depends on a ring of actin and myosin that contracts in late mitosis at a site midway between the segregated chromosomes. In animal cells, the positioning of the contractile ring is determined by signals emanating from the microtubules of the anaphase spindle. Dephosphorylation of Cdk targets, which results from Cdk inactivation in anaphase, triggers cytokinesis at the correct time after anaphase. After cytokinesis, the cell enters a stable G_1 state of low Cdk activity, where it awaits signals to enter a new cell cycle.

CONTROL OF CELL DIVISION AND CELL GROWTH

A fertilized mouse egg and a fertilized human egg are similar in size, yet they produce animals of very different sizes. What factors in the control of cell behavior in humans and mice are responsible for these size differences? The same fundamental question can be asked for each organ and tissue in an animal's body. What factors in the control of cell behavior explain the length of an elephant's trunk or the size of its brain or its liver? These questions are largely unanswered, at least in part because they have received relatively little attention compared with other questions in cell and developmental biology. It is nevertheless possible to say what the ingredients of an answer must be.

The size of an organ or organism depends mainly on its total cell mass, which depends on both the total number of cells and the size of the cells. Cell number, in turn, depends on the amounts of cell division and cell death. Organ and body size are therefore determined by three fundamental processes: cell growth, cell division, and cell death. Each is tightly regulated—both by intracellular programs and by extracellular signal molecules that control these programs.

The extracellular signal molecules that regulate cell size and cell number are generally soluble secreted proteins, proteins bound to the surface of cells, or components of the extracellular matrix. They can be divided operationally into three major classes:

1. *Mitogens*, which stimulate cell division, primarily by triggering a wave of G_1/S-Cdk activity that relieves intracellular negative controls that otherwise block progress through the cell cycle.
2. *Growth factors*, which stimulate cell growth (an increase in cell mass) by promoting the synthesis of proteins and other macromolecules and by inhibiting their degradation.
3. *Survival factors*, which promote cell survival by suppressing the form of programmed cell death known as *apoptosis*.

Many extracellular signal molecules promote all of these processes, while others promote one or two of them. Indeed, the term *growth factor* is often used inappropriately to describe a factor that has any of these activities. Even worse, the term *cell growth* is often used to mean an increase in cell number, or *cell proliferation*.

In addition to these three classes of stimulating signals, there are extracellular signal molecules that suppress cell proliferation, cell growth, or both; in general, less is known about them. There are also extracellular signal molecules that activate apoptosis.

In this section, we focus primarily on how mitogens and other factors, such as DNA damage, control the rate of cell division. We then turn to the important but poorly understood problem of how a proliferating cell coordinates its growth with cell division so as to maintain its appropriate size. We discuss the control of cell survival and cell death by apoptosis in Chapter 18.

Mitogens Stimulate Cell Division

Unicellular organisms tend to grow and divide as fast as they can, and their rate of proliferation depends largely on the availability of nutrients in the environment. The cells of a multicellular organism, however, divide only when the organism needs more cells. Thus, for an animal cell to proliferate, it must receive stimulatory extracellular signals, in the form of **mitogens**, from other cells, usually its neighbors. Mitogens overcome intracellular braking mechanisms that block progress through the cell cycle.

One of the first mitogens to be identified was *platelet-derived growth factor (PDGF)*, and it is typical of many others discovered since. The path to its isolation began with the observation that fibroblasts in a culture dish proliferate when provided with *serum* but not when provided with *plasma*. Plasma is prepared by removing the cells from blood without allowing clotting to occur; serum is prepared by allowing blood to clot and taking the cell-free liquid that remains. When blood clots, platelets incorporated in the clot are stimulated to release the contents of their secretory vesicles (**Figure 17–61**). The superior ability of serum to support cell proliferation suggested that platelets contain one or more mitogens. This hypothesis was confirmed by showing that extracts of platelets could serve instead of serum to stimulate fibroblast proliferation. The crucial factor in the extracts was shown to be a protein, which was subsequently purified and named PDGF. In the body, PDGF liberated from blood clots helps stimulate cell division during wound healing.

PDGF is only one of over 50 proteins that are known to act as mitogens. Most of these proteins have a broad specificity. PDGF, for example, can stimulate many types of cells to divide, including fibroblasts, smooth muscle cells, and

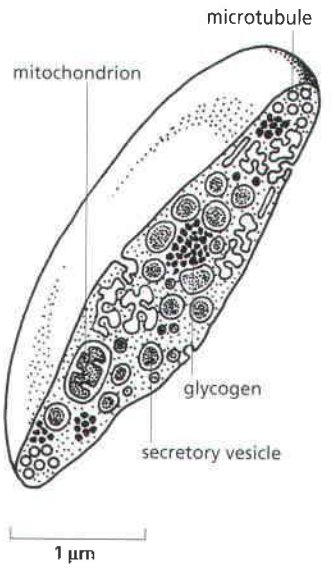


Figure 17–61 **A platelet.** Platelets are miniature cells without a nucleus. They circulate in the blood and help stimulate blood clotting at sites of tissue damage, thereby preventing excessive bleeding. They also release various factors that stimulate healing. The platelet shown here has been cut in half to show its secretory vesicles, some of which contain platelet-derived growth factor (PDGF).

neuroglial cells. Similarly, *epidermal growth factor* (*EGF*) acts not only on epidermal cells but also on many other cell types, including both epithelial and nonepithelial cells. Some mitogens, however, have a narrow specificity; *erythropoietin*, for example, only induces the proliferation of red blood cell precursors. Many mitogens, including PDGF, also have other actions beside the stimulation of cell division: they can stimulate cell growth, survival, differentiation, or migration, depending on the circumstances and the cell type.

In some tissues, inhibitory extracellular signal proteins oppose the positive regulators and thereby inhibit organ growth. The best-understood inhibitory signal proteins are TGFβ and its relatives. TGFβ inhibits the proliferation of several cell types, either by blocking cell-cycle progression in G_1 or by stimulating apoptosis.

Cells Can Delay Division by Entering a Specialized Nondividing State

In the absence of a mitogenic signal to proliferate, Cdk inhibition in G_1 is maintained by the multiple mechanisms discussed earlier, and progression into a new cell cycle is blocked. In some cases, cells partly disassemble their cell-cycle control system and exit from the cycle to a specialized, nondividing state called **G_0**.

Most cells in our body are in G_0, but the molecular basis and reversibility of this state vary in different cell types. Most of our neurons and skeletal muscle cells, for example, are in a *terminally differentiated* G_0 state, in which their cell-cycle control system is completely dismantled: the expression of the genes encoding various Cdks and cyclins are permanently turned off, and cell division rarely occurs. Other cell types withdraw from the cell cycle only transiently and retain the ability to reassemble the cell-cycle control system quickly and reenter the cycle. Most liver cells, for example, are in G_0, but they can be stimulated to divide if the liver is damaged. Still other types of cells, including fibroblasts and some lymphocytes, withdraw from and re-enter the cell cycle repeatedly throughout their lifetime.

Almost all the variation in cell-cycle length in the adult body occurs during the time the cell spends in G_1 or G_0. By contrast, the time a cell takes to progress from the beginning of S phase through mitosis is usually brief (typically 12–24 hours in mammals) and relatively constant, regardless of the interval from one division to the next.

Mitogens Stimulate G_1-Cdk and G_1/S-Cdk Activities

For the vast majority of animal cells, mitogens control the rate of cell division by acting in the G_1 phase of the cell cycle. As discussed earlier, multiple mechanisms act during G_1 to suppress Cdk activity and thereby block entry into S phase. Mitogens release these brakes on Cdk activity, thereby allowing S phase to begin.

As we discuss in Chapter 15, mitogens interact with cell-surface receptors to trigger multiple intracellular signaling pathways. One major pathway acts through the small GTPase **Ras**, which leads to the activation of a *MAP kinase cascade*. This leads to an increase in the production of gene regulatory proteins, including **Myc**. Myc is thought to promote cell-cycle entry by several mechanisms, one of which is to increase the expression of genes encoding G_1 cyclins (D cyclins), thereby increasing G_1-Cdk (cyclin D–Cdk4) activity. As we discuss later, Myc also has a major role in stimulating the transcription of genes that increase cell growth.

The key function of G_1-Cdk complexes in animal cells is to activate a group of gene regulatory factors called the **E2F proteins**, which bind to specific DNA sequences in the promoters of a wide variety of genes that encode proteins required for S-phase entry, including G_1/S-cyclins, S-cyclins, and proteins involved in DNA synthesis and chromosome duplication. In the absence of mitogenic stimulation, E2F-dependent gene expression is inhibited by an interaction between E2F and members of the **retinoblastoma protein** (**Rb**) family.

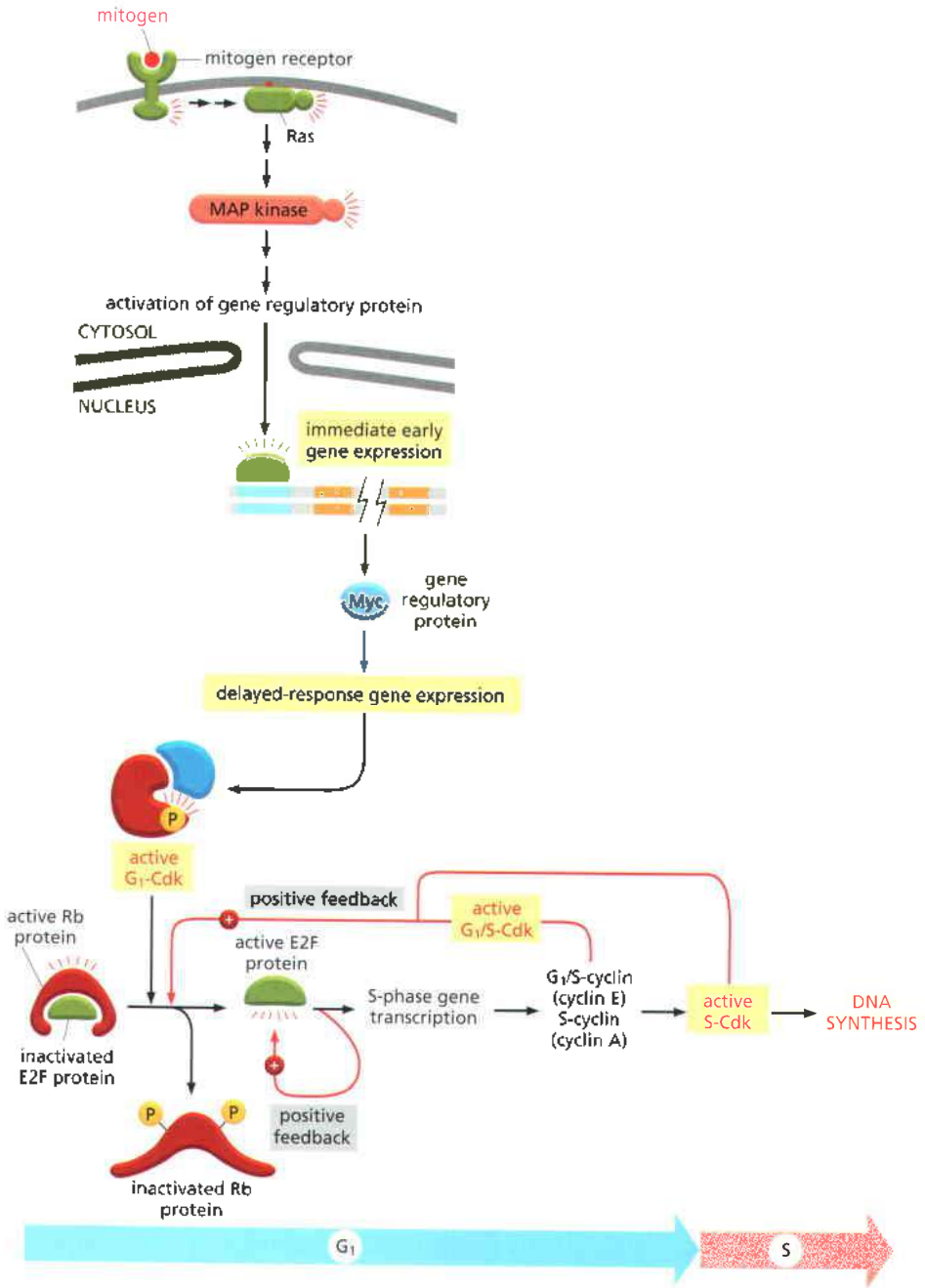


Figure 17–62 Mechanisms controlling cell-cycle entry and S-phase initiation in animal cells. As discussed in Chapter 15, mitogens bind to cell-surface receptors to initiate intracellular signaling pathways. One of the major pathways involves activation of the small GTPase Ras, which activates a MAP kinase cascade, leading to increased expression of numerous *immediate early* genes, including the gene encoding the gene regulatory protein Myc. Myc increases the expression of many *delayed-response* genes, including some that lead to increased G_1-Cdk activity (cyclin D–Cdk4), which triggers the phosphorylation of members of the Rb family of proteins. This inactivates the Rb proteins, freeing the gene regulatory protein E2F to activate the transcription of G_1/S genes, including the genes for a G_1/S-cyclin (cyclin E) and S-cyclin (cyclin A). The resulting G_1/S-Cdk and S-Cdk activities further enhance Rb protein phosphorylation, forming a positive feedback loop. E2F proteins also stimulate the transcription of their own genes, forming another positive feedback loop.

When cells are stimulated to divide by mitogens, active G_1-Cdk accumulates and phosphorylates Rb family members, reducing their binding to E2F. The liberated E2F proteins then activate expression of their target genes (Figure 17–62).

This transcriptional control system, like so many other control systems that regulate the cell cycle, includes feedback loops that sharpen the G_1/S transition. The liberated E2F proteins, for example, increase the transcription of their own genes. In addition, E2F-dependent transcription of G_1/S-cyclin (cyclin E) and S-cyclin (cyclin A) genes leads to increased G_1/S-Cdk and S-Cdk activities, which in turn increase Rb protein phosphorylation and promote further E2F release (see Figure 17–62).

The central member of the Rb family, the Rb protein itself, was identified originally through studies of an inherited form of eye cancer in children, known as *retinoblastoma* (discussed in Chapter 20). The loss of both copies of the Rb gene leads to excessive cell proliferation in the developing retina, suggesting that the Rb protein is particularly important for restraining cell division in this tissue. The complete loss of Rb does not immediately cause increased proliferation of

retinal or other cell types, in part because Cdh1 and CKIs also help inhibit progression through G_1 and in part because other cell types contain Rb-related proteins that provide backup support in the absence of Rb. It is also likely that other proteins, unrelated to Rb, help to regulate the activity of E2F.

Additional layers of control promote an overwhelming increase in S-Cdk activity at the beginning of S phase. We mentioned earlier that the APC/C activator Cdh1 suppresses cyclin levels after mitosis. In animal cells, however, G_1- and G_1/S-cyclins are resistant to Cdh1 and can therefore act unopposed by the APC/C to promote Rb protein phosphorylation and E2F-dependent gene expression. S-cyclin, by contrast, is not resistant to Cdh1, and its level is initially restrained by Cdh1-APC/C activity. However, G_1/S-Cdk also phosphorylates and inactivates Cdh1-APC/C, thereby allowing the accumulation of S-cyclin, further promoting S-Cdk activation. G_1/S-Cdk also inactivates CKI proteins that suppress S-Cdk activity. The overall effect of all these interactions is the rapid and complete activation of the S-Cdk complexes required for S-phase initiation.

DNA Damage Blocks Cell Division: The DNA Damage Response

Progression through the cell cycle, and thus the rate of cell proliferation, is controlled not only by extracellular mitogens but also by other extracellular and intracellular mechanisms. One of the most important influences is DNA damage, which can occur as a result of spontaneous chemical reactions in DNA, errors in DNA replication, or exposure to radiation or certain chemicals. It is essential that the cell repair damaged chromosomes before attempting to duplicate or segregate them. The cell-cycle control system can readily detect DNA damage and arrest the cycle at either of two checkpoints—one at Start in late G_1, which prevents entry into the cell cycle and into S phase, and one at the G_2/M checkpoint, which prevents entry into mitosis (see Figure 17–21).

DNA damage initiates a signaling pathway by activating one of a pair of related protein kinases called **ATM** and **ATR**, which associate with the site of damage and phosphorylate various target proteins, including two other protein kinases called *Chk1* and *Chk2*. Together these various kinases phosphorylate other target proteins that lead to cell-cycle arrest. A major target is the gene regulatory protein **p53**, which stimulates transcription of the gene encoding a CKI protein called *p21;* this protein binds to G_1/S-Cdk and S-Cdk complexes and inhibits their activities, thereby helping to block entry into the cell cycle (**Figure 17–63**). <TGAA>

DNA damage activates p53 by an indirect mechanism. In undamaged cells, p53 is highly unstable and is present at very low concentrations. This is largely because it interacts with another protein, *Mdm2*, which acts as a ubiquitin ligase that targets p53 for destruction by proteasomes. Phosphorylation of p53 after DNA damage reduces its binding to Mdm2. This decreases p53 degradation, which results in a marked increase in p53 concentration in the cell. In addition, the decreased binding to Mdm2 enhances the ability of p53 to stimulate gene transcription (see Figure 17–63).

The protein kinases Chk1 and Chk2 also block cell cycle progression by phosphorylating members of the Cdc25 family of protein phosphatases, thereby inhibiting their function. As described earlier, these kinases are particularly important in the activation of M-Cdk at the beginning of mitosis (see Figure 17–25). Thus, the inhibition of Cdc25 activity by DNA damage helps block entry into mitosis (see Figure 17–21).

The DNA-damage response also detects problems that arise when a replication fork fails during DNA replication. When nucleotides are depleted, for example, replication forks stall during the elongation phase of DNA synthesis. To prevent the cell from attempting to segregate partially replicated chromosomes, the same mechanisms that respond to DNA damage detect the stalled replication forks and block entry into mitosis until the problems at the replication fork are resolved.

The DNA damage response is not essential for normal cell division if environmental conditions are ideal. Conditions are rarely ideal, however: a low level

of DNA damage occurs in the normal life of any cell, and this damage accumulates in the cell's progeny if the damage response is not functioning. Over the long term, the accumulation of genetic damage in cells lacking the DNA damage response leads to an increased frequency of cancer-promoting mutations. Indeed, mutations in the *p53* gene occur in at least half of all human cancers (discussed in Chapter 20). This loss of p53 function allows the cancer cell to accumulate mutations more readily. Similarly, a rare genetic disease known as *ataxia telangiectasia* is caused by a defect in ATM, one of the protein kinases that is activated in response to x-ray-induced DNA damage; patients with this disease are very sensitive to x-rays and suffer from increased rates of cancer.

What happens if DNA damage is so severe that repair is not possible? The answer differs in different organisms. Unicellular organisms such as budding yeast transiently arrest their cell cycle to try to repair the damage, but the cycle resumes even if the repair cannot be completed. For a single-celled organism, life with mutations is apparently better than no life at all. In multicellular organisms, however, the health of the organism takes precedence over the life of an individual cell. Cells that divide with severe DNA damage threaten the life of the organism, since genetic damage can often lead to cancer and other diseases. Thus, animal cells with severe DNA damage do not attempt to continue division, but instead commit suicide by undergoing apoptosis. Thus, unless the DNA damage is repaired, the DNA damage response can lead to either cell-cycle

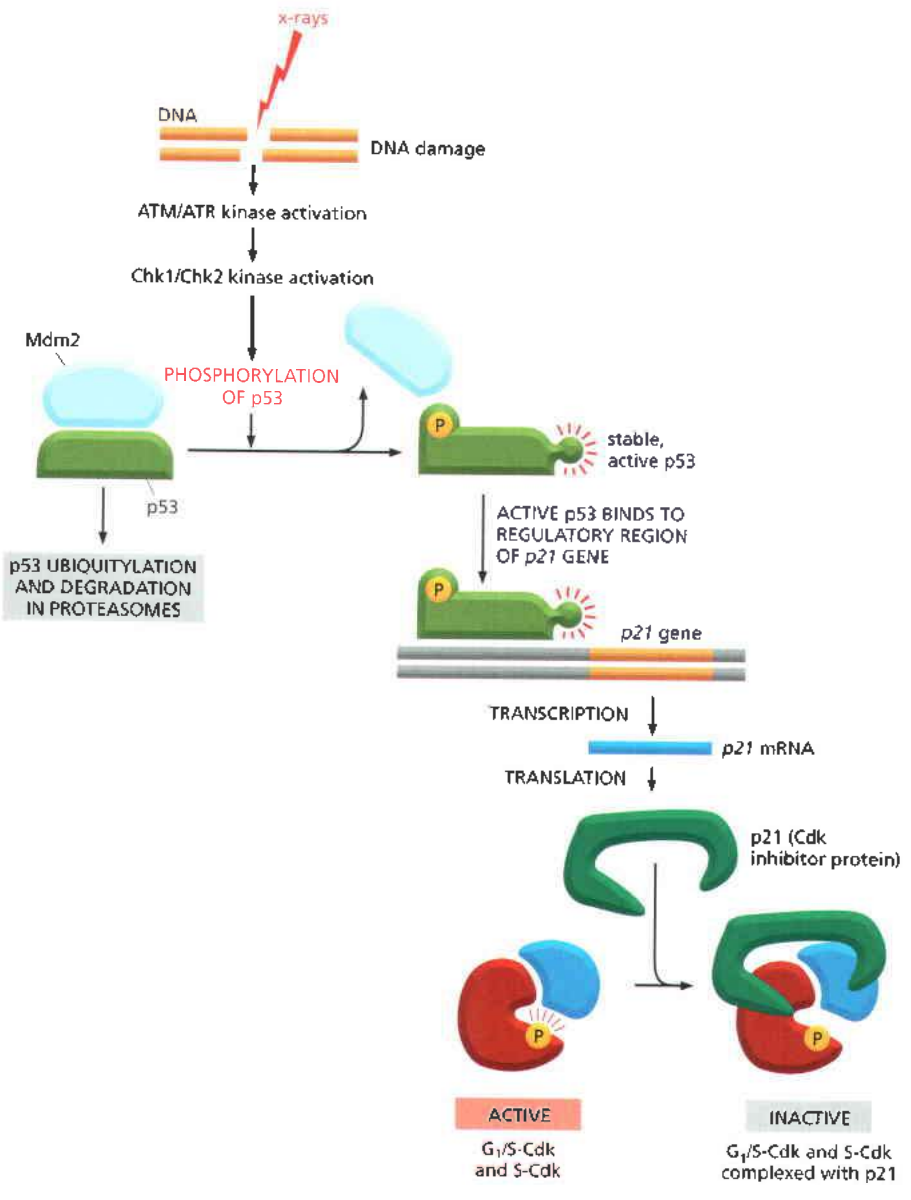


Figure 17–63 How DNA damage arrests the cell cycle in G_1. When DNA is damaged, various protein kinases are recruited to the site of damage and initiate a signaling pathway that causes cell-cycle arrest. The first kinase at the damage site is either ATM or ATR, depending on the type of damage. Additional protein kinases, called Chk1 and Chk2, are then recruited and activated, resulting in the phosphorylation of the gene regulatory protein p53. Mdm2 normally binds to p53 and promotes its ubiquitylation and destruction in proteasomes. Phosphorylation of p53 blocks its binding to Mdm2; as a result, p53 accumulates to high levels and stimulates transcription of the gene that encodes the CKI protein p21. The p21 binds and inactivates G_1/S-Cdk and S-Cdk complexes, arresting the cell in G_1. In some cases, DNA damage also induces either the phosphorylation of Mdm2 or a decrease in Mdm2 production, which causes a further increase in p53 (not shown).

arrest or cell death. As we discuss in the next chapter, DNA damage-induced apoptosis often depends on the activation of p53. Indeed, it is this apoptosis-promoting function of p53 that is apparently most important in protecting us against cancer.

Many Human Cells Have a Built-In Limitation on the Number of Times They Can Divide

Many human cells divide a limited number of times before they stop and undergo a permanent cell-cycle arrest. Fibroblasts taken from normal human tissue, for example, go through only about 25–50 population doublings when cultured in a standard mitogenic medium. Toward the end of this time, proliferation slows down and finally halts, and the cells enter a nondividing state from which they never recover. This phenomenon is called **replicative cell senescence**, although it is unlikely to be responsible for the senescence (aging) of the organism. Organism senescence is thought to depend, in part, on progressive oxidative damage to long-lived macromolecules, as strategies that reduce metabolism (such as reduced food intake), and thereby reduce the production of reactive oxygen species, can extend the lifespan of experimental animals.

Replicative cell senescence in human fibroblasts seems to be caused by changes in the structure of the **telomeres**, the repetitive DNA sequences and associated proteins at the ends of chromosomes. As discussed in Chapter 5, when a cell divides, telomeric DNA sequences are not replicated in the same manner as the rest of the genome but instead are synthesized by the enzyme **telomerase**. Telomerase also promotes the formation of protein cap structures that protect the chromosome ends. Because human fibroblasts, and many other human somatic cells, are deficient in telomerase, their telomeres become shorter with every cell division, and their protective protein caps progressively deteriorate. Eventually, the exposed chromosome ends are sensed as DNA damage, which activates a p53-dependent cell-cycle arrest that resembles the arrest caused by other types of DNA damage (see Figure 17–63). Rodent cells, by contrast, maintain telomerase activity when they proliferate in culture and therefore do not have such a telomere-dependent mechanism for limiting proliferation. The forced expression of telomerase in normal human fibroblasts, using genetic engineering techniques, blocks this form of senescence. Unfortunately, most cancer cells have regained the ability to produce telomerase and therefore maintain telomere function as they proliferate; as a result, they do not undergo replicative cell senescence.

Abnormal Proliferation Signals Cause Cell-Cycle Arrest or Apoptosis, Except in Cancer Cells

Many of the components of mitogenic signaling pathways are encoded by genes that were originally identified as cancer-promoting genes, or *oncogenes*, because mutations in them contribute to the development of cancer. The mutation of a single amino acid in the small GTPase Ras, for example, causes the protein to become permanently overactive, leading to constant stimulation of Ras-dependent signaling pathways, even in the absence of mitogenic stimulation. Similarly, mutations that cause an overexpression of Myc stimulate excessive cell growth and proliferation and thereby promote the development of cancer.

Surprisingly, however, when a hyperactivated form of Ras or Myc is experimentally overproduced in most normal cells, the result is not excessive proliferation but the opposite: the cells undergo either cell-cycle arrest or apoptosis. The normal cell seems able to detect abnormal mitogenic stimulation, and it responds by preventing further division. Such responses help prevent the survival and proliferation of cells with various cancer-promoting mutations.

Although it is not known how a cell detects excessive mitogenic stimulation, such stimulation often leads to the production of a cell-cycle inhibitor protein called *Arf*, which binds and inhibits Mdm2. As discussed earlier, Mdm2 normally

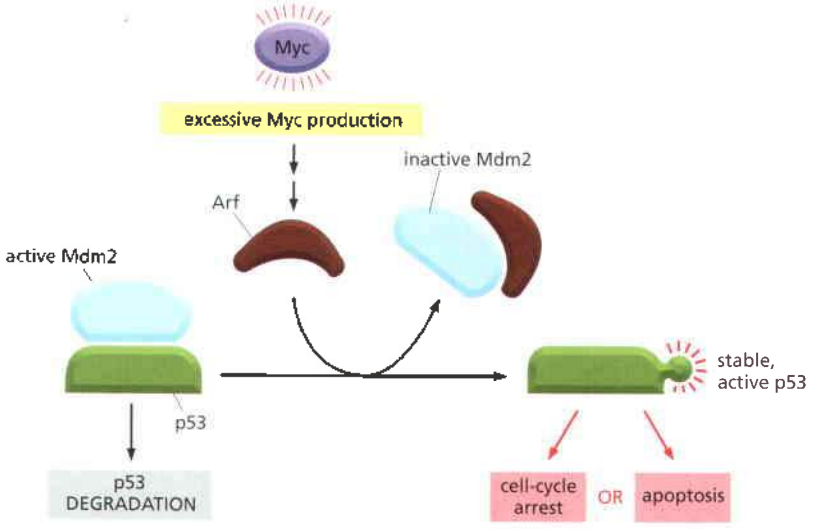


Figure 17–64 Cell-cycle arrest or apoptosis induced by excessive stimulation of mitogenic pathways. Abnormally high levels of Myc cause the activation of Arf, which binds and inhibits Mdm2 and thereby increases p53 levels (see Figure 17–60). Depending on the cell type and extracellular conditions, p53 then causes either cell-cycle arrest or apoptosis.

promotes p53 degradation. Activation of Arf therefore causes p53 levels to increase, inducing either cell-cycle arrest or apoptosis (**Figure 17–64**).

How do cancer cells ever arise if these mechanisms block the division or survival of mutant cells with overactive proliferation signals? The answer is that the protective system is often inactivated in cancer cells by mutations in the genes that encode essential components of the checkpoint responses, such as Arf or p53 or the proteins that help activate them.

Organism and Organ Growth Depend on Cell Growth

For an organism or organ to grow, cell division is not enough. If cells proliferated without growing, they would get progressively smaller and there would be no net increase in total cell mass. In most proliferating cell populations, therefore, cell growth accompanies cell division. In single-celled organisms such as yeasts, both cell growth and cell division require only nutrients. In animals, by contrast, both cell growth and cell proliferation depend on extracellular signal molecules, produced by other cells, which we call **growth factors** and mitogens, respectively.

Like mitogens, the extracellular growth factors that stimulate animal cell growth bind to receptors on the cell surface and activate intracellular signaling pathways. These pathways stimulate the accumulation of proteins and other macromolecules, and they do so by both increasing their rate of synthesis and decreasing their rate of degradation. They also trigger increased uptake of nutrients and production of the ATP required to fuel increased protein synthesis. One of the most important intracellular signaling pathways activated by growth factor receptors involves the enzyme *PI 3-kinase*, which adds a phosphate from ATP to the 3 position of inositol phospholipids in the plasma membrane. As discussed in Chapter 15, the activation of PI 3-kinase leads to the activation of a kinase called *TOR*, which lies at the heart of growth regulatory pathways in all eucaryotes. TOR activates many targets in the cell that stimulate metabolic processes and increase protein synthesis. One target is a protein kinase called *S6 kinase (S6K)*, which phosphorylates ribosomal protein S6, increasing the ability of ribosomes to translate a subset of mRNAs that mostly encode ribosomal components. TOR also indirectly activates a translation initiation factor called *eIF4E* and directly activates gene regulatory proteins that promote the increased expression of genes encoding ribosomal subunits (**Figure 17–65**).

Proliferating Cells Usually Coordinate Their Growth and Division

For proliferating cells to maintain a constant size, they must coordinate their growth with cell division to ensure that cell size doubles with each division: if cells grow too slowly, they will get smaller with each division, and if they grow

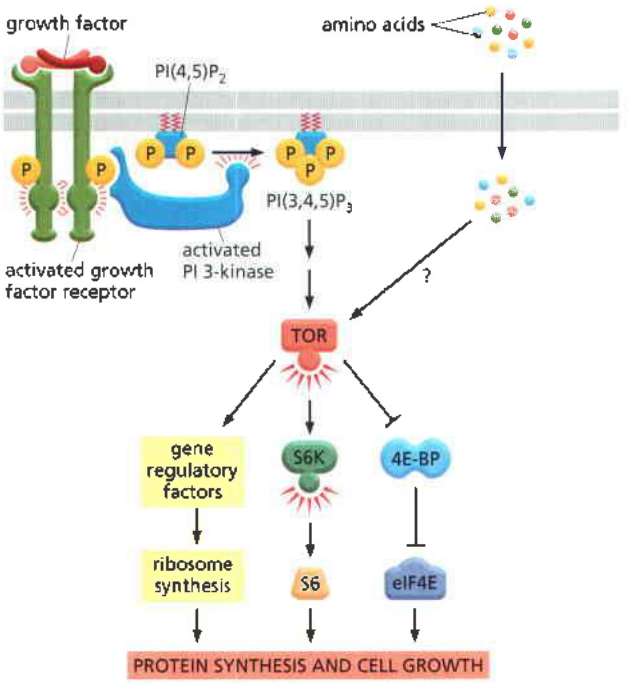


Figure 17–65 Stimulation of cell growth by extracellular growth factors and nutrients. As discussed in Chapter 15, the occupation of cell-surface receptors by growth factors leads to the activation of PI 3-kinase, which promotes protein synthesis through a complex signaling pathway that leads to the activation of the protein kinase TOR; extracellular nutrients such as amino acids also help activate TOR by an unknown pathway. TOR employs multiple mechanisms to stimulate protein synthesis, as shown; it also inhibits protein degradation (not shown). Growth factors also stimulate increased production of the gene regulatory protein Myc (not shown), which activates the transcription of various genes that promote cell metabolism and growth. 4E-BP is an inhibitor of the translation initiation factor eIF4E.

too fast, they will get larger with each division. It is not clear how cells achieve this coordination, but it is likely to involve multiple mechanisms that vary in different organisms and even in different cell types of the same organism (**Figure 17–66**).

Animal cell growth and division are not always coordinated, however. In many cases, they are completely uncoupled to allow growth without division or division without growth. Muscle cells and nerve cells, for example, can grow dramatically after they have permanently withdrawn from the cell cycle. Similarly, the eggs of many animals grow to an extremely large size without dividing; after fertilization, however, this relationship is reversed, and many rounds of division occur without growth (see Figure 17–9).

Compared to cell division, there has been surprisingly little study of how cell size is controlled in animals. As a result, it remains a mystery how cell size is determined and why different cell types in the same animal grow to be so different in size (**Figure 17–67**). One of the best-understood cases in mammals is the adult *sympathetic neuron*, which has permanently withdrawn from the cell cycle. Its size depends on the amount of *nerve growth factor* (NGF) secreted by the target cells it innervates; the greater the amount of NGF the neuron has access to, the larger it becomes. It seems likely that the genes a cell expresses set limits on the size it can be, while extracellular signal molecules and nutrients

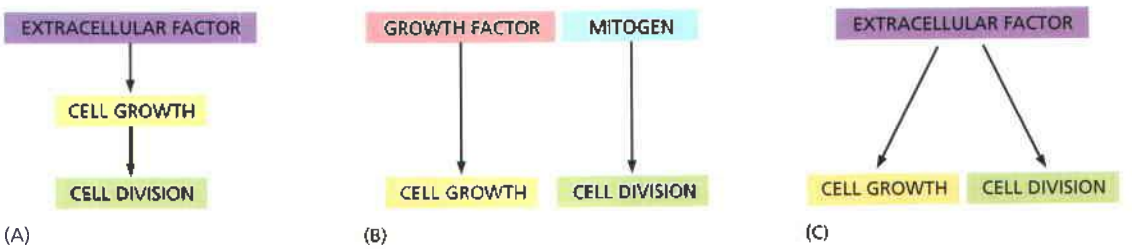


Figure 17–66 Potential mechanisms for coordinating cell growth and division. In proliferating cells, cell size is maintained by mechanisms that coordinate rates of cell division and cell growth. Numerous alternative coupling mechanisms are thought to exist, and different cell types appear to employ different combinations of these mechanisms. (A) In many cell types—particularly yeast—the rate of cell division is governed by the rate of cell growth, so that division occurs only when growth rate achieves some minimal threshold; in yeasts, it is mainly the levels of extracellular nutrients that regulate the rate of cell growth and thereby the rate of cell division. (B) In some animal cell types, growth and division can each be controlled by separate extracellular factors (growth factors and mitogens, respectively), and cell size depends on the relative levels of the two types of factors. (C) Some extracellular factors can stimulate both cell growth and cell division by simultaneously activating signaling pathways that promote growth and other pathways that promote cell-cycle progression.

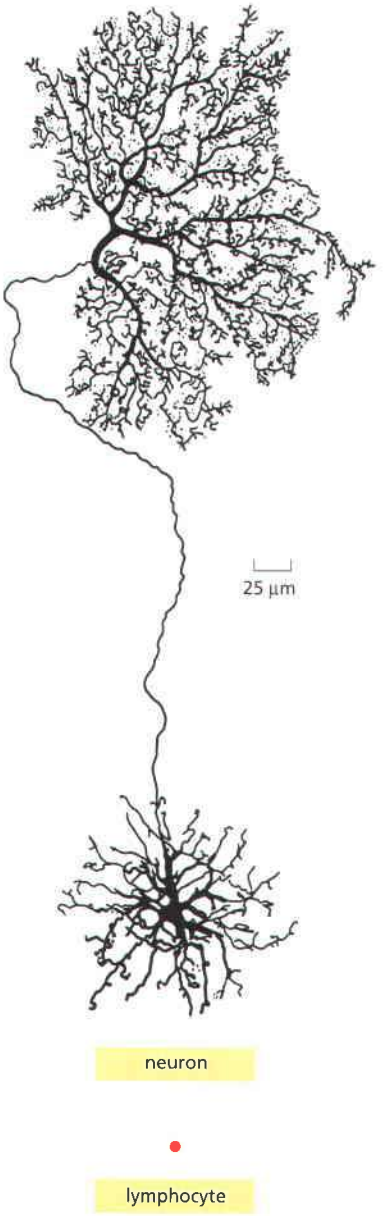


Figure 17–67 **The size difference between a neuron (from the retina) and a lymphocyte in a mammal.** Both cells contain the same amount of DNA. A neuron grows progressively larger after it has permanently withdrawn from the cell cycle. During this time, the ratio of cytoplasm to DNA increases enormously (by a factor of more than 10^5 for some neurons). (Neuron from B.B. Boycott, in Essays on the Nervous System [R. Bellairs and E.G. Gray, eds]. Oxford, UK: Clarendon Press, 1974.)

regulate the size within these limits. The challenge is to identify the relevant genes and signal molecules for each cell type.

Neighboring Cells Compete for Extracellular Signal Proteins

When most types of mammalian cells are cultured in a dish in the presence of serum, they adhere to the bottom of the dish, spread out, and divide until they form a confluent monolayer. Each cell is attached to the dish and contacts its neighbors on all sides. At this point, normal cells, unlike cancer cells, stop proliferating—a phenomenon known as *density-dependent inhibition of cell division.* This phenomenon was originally described in terms of "contact inhibition" of cell division, but it is unlikely that cell–cell contact interactions are solely responsible. The cell population density at which cell proliferation ceases in the confluent monolayer increases as the concentration of serum in the medium increases. Moreover, if a stream of fresh culture medium is passed over a confluent layer of fibroblasts to increase the supply of mitogens, the cells under the stream are induced to divide (**Figure 17–68**). Thus, density-dependent inhibition of cell proliferation seems to reflect, in part at least, the ability of a cell to deplete the medium around itself of extracellular mitogens, thereby depriving its neighbors.

This type of competition could be important for cells in tissues as well as in culture, because it prevents them from proliferating beyond a certain population density, determined by the available amounts of mitogens, growth factors, and survival factors. The amounts of these factors in tissues are usually limiting, in that increasing their amounts results in an increase in cell number, cell size, or both. Thus, the amounts of these factors in tissues have important roles in determining cell size and number, and possibly the final size of the organ or tissue.

The overall size of a tissue may also be governed in some cases by extracellular inhibitory factors. *Myostatin,* for example, is a TGFβ family member that normally inhibits the proliferation of myoblasts that fuse to form skeletal muscle cells. When the gene that encodes myostatin is deleted in mice, muscles grow to be several times larger than normal. Remarkably, two breeds of cattle that were bred for large muscles have both turned out to have mutations in the gene encoding myostatin (**Figure 17–69**).

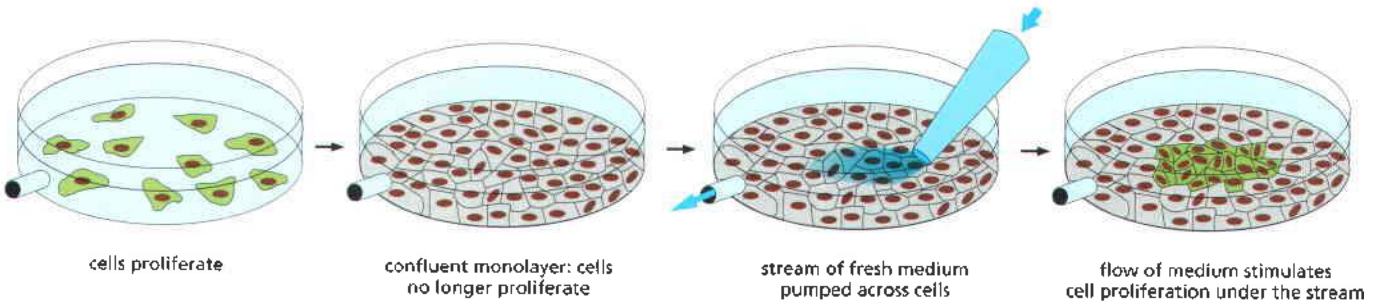


Figure 17–68 **The effect of fresh medium on a confluent cell monolayer.** Cells in a confluent monolayer do not divide *(gray).* The cells resume dividing *(green)* when exposed directly to fresh culture medium. Apparently, in the undisturbed confluent monolayer, proliferation has halted because the medium close to the cells is depleted of mitogens, for which the cells compete.



Figure 17–69 The effects of a myostatin mutation on muscle size. The mutation leads to a great increase in the mass of muscle tissue, as illustrated in this Belgian Blue bull. The Belgian Blue was produced by cattle breeders and was only recently found to have a mutation in the *Myostatin* gene. (From H.L. Sweeney, *Sci. Am.* 291:62, 2004. With permission from Scientific American.)

Animals Control Total Cell Mass by Unknown Mechanisms

The size of an animal or one of its organs depends largely on the number and size of the cells it contains—that is, on total cell mass. Remarkably, animals can somehow assess the total cell mass in a tissue or organ and regulate it: in many circumstances, for example, if cell size is experimentally increased or decreased in an organ, cell numbers adjust to maintain a normal organ size. This has been most dramatically illustrated by experiments in salamanders, in which cell size was manipulated by altering cell ploidy (in all organisms, the size of a cell is proportional to its ploidy, or genome content). Salamanders of different ploidies are the same size but have different numbers of cells. Individual cells in a pentaploid salamander are about five times the volume of those in a haploid salamander, and in each organ the pentaploids have only one-fifth as many cells as their haploid cousins, so that the organs are about the same size in the two animals (**Figure 17–70** and **Figure 17–71**). Evidently, in this case (and in many others) the size of organs and organisms depends on mechanisms that can somehow measure total cell mass. How animals measure and adjust total mass remains a mystery, however.

The development of limbs and organs of specific size and shape depends on complex positional controls, as well as on local concentrations of extracellular signal proteins that stimulate or inhibit cell growth, division, and survival. As we discuss in Chapter 22, we now know many of the genes that help pattern these processes in the embryo. A great deal remains to be learned, however, about how these genes regulate cell growth, division, survival, and differentiation to generate a complex organism.

The controls that govern these processes in an adult body are also poorly understood. When a skin wound heals in a vertebrate, for example, about a dozen cell types, ranging from fibroblasts to Schwann cells, must be regenerated in appropriate numbers, sizes, and positions to reconstruct the lost tissue. The

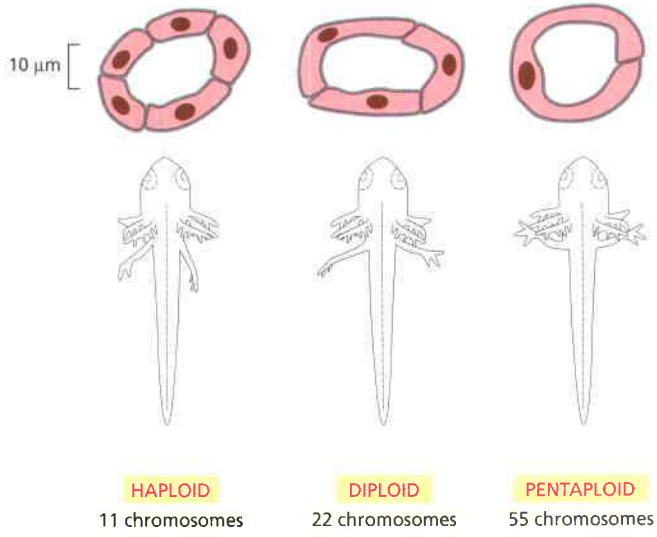


Figure 17–70 Sections of kidney tubules from salamander larvae of different ploidies. In all organisms, from bacteria to humans, cell size is proportional to ploidy. Pentaploid salamanders, for example, have cells that are much larger than those of haploid salamanders. The animals and their individual organs, however, are the same size because each tissue in the pentaploid animal contains fewer cells. This indicates that the size of an organism or organ is not controlled simply by counting cell divisions or cell numbers; total cell mass must somehow be regulated. (Adapted from G. Fankhauser, in Analysis of Development [B.H. Willier, P.A. Weiss, and V. Hamburger, eds.], pp. 126–150. Philadelphia: Saunders, 1955.)

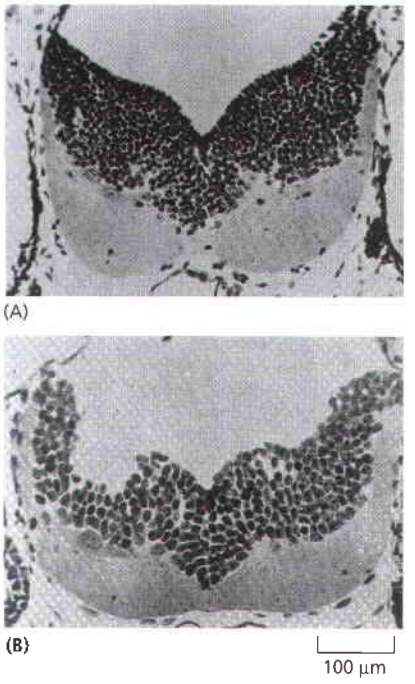


Figure 17–71 **The hindbrain in a haploid and in a tetraploid salamander.** (A) This light micrograph shows a cross section of the hindbrain of a haploid salamander. (B) A corresponding cross section of the hindbrain of a tetraploid salamander, revealing how reduced cell numbers compensate for increased cell size, so that the overall size of the hindbrain is the same in the two animals. (From G. Fankhauser, *Int. Rev. Cytol.* 1:165–193, 1952. With permission from Elsevier.)

mechanisms that control cell growth and proliferation in tissues are likewise central to understanding cancer, a disease in which the controls go wrong, as discussed in Chapter 20.

Summary

In multicellular animals, cell size, cell division, and cell death are carefully controlled to ensure that the organism and its organs achieve and maintain an appropriate size. Mitogens stimulate the rate of cell division by removing intracellular molecular brakes that restrain cell-cycle progression in G_1. Growth factors promote cell growth (an increase in cell mass) by stimulating the synthesis and inhibiting the degradation of macromolecules. For proliferating cells to maintain a constant cell size, they employ multiple mechanisms to ensure that cell growth is coordinated with cell division. Animals maintain the normal size of their tissues and organs by adjusting cell size to compensate for changes in cell number, or vice versa. The mechanisms that make this possible are not known.

PROBLEMS

Which statements are true? Explain why or why not.

17–1 Since there are about 10^{13} cells in an adult human, and about 10^{10} cells die and are replaced each day, we become new people every three years.

17–2 The regulation of cyclin–Cdk complexes depends entirely on phosphorylation and dephosphorylation.

17–3 In order for proliferating cells to maintain a relatively constant size, the length of the cell cycle must match the time it takes for the cell to double in size.

17–4 While other proteins come and go during the cell cycle, the proteins of the origin recognition complex remain bound to the DNA throughout.

17–5 Chromosomes are positioned on the metaphase plate by equal and opposite forces that pull them toward the two poles of the spindle.

17–6 If we could turn on telomerase activity in all our cells, we could prevent aging.

Discuss the following problems.

17–7 Many cell-cycle genes from human cells function perfectly well when expressed in yeast cells. Why do you suppose that is considered remarkable? After all, many human genes encoding enzymes for metabolic reactions also function in yeast, and no one thinks that is remarkable.

17–8 You have isolated a new *Cdc* mutant of budding yeast that forms colonies at 25°C but not at 37°C. You would now like to isolate the wild-type gene that corresponds to the defective gene in your *Cdc* mutant. How might you isolate the wild-type gene using a plasmid-based DNA library prepared from wild-type yeast cells?

17–9 You have isolated a temperature-sensitive mutant of budding yeast. It proliferates well at 25°C, but at 35°C all the cells develop a large bud and then halt their progression through the cell cycle. The characteristic morphology of the cells at the time they stop cycling is known as the landmark morphology.

It is very difficult to obtain synchronous cultures of this yeast, but you would like to know exactly where in the cell cycle the temperature-sensitive gene product must function—its execution point, in the terminology of the field—in order for the cell to complete the cycle. A clever friend, who has a good microscope with a heated stage and a video camera, suggests that you take movies of a field of cells as they experience the temperature increase, and follow the morphology of the cells as they stop cycling. Since the cells do not move much, it is relatively simple to study individual cells. To make sense of what you see, you arrange a circle of pictures of cells at the start of the experiment in order of the size of their daughter buds. You then find the corresponding pictures of those same cells 6 hours later, when growth and division has completely stopped. The results with your mutant are shown in Figure Q17–1.

A. Indicate on the diagram in Figure Q17–1 where the execution point for your mutant lies.

B. Does the execution point correspond to the time at which the cell cycle is arrested in your mutant? How can you tell?

17–10 The yeast cohesin subunit Scc1, which is essential for sister-chromatid pairing, can be artificially regulated for

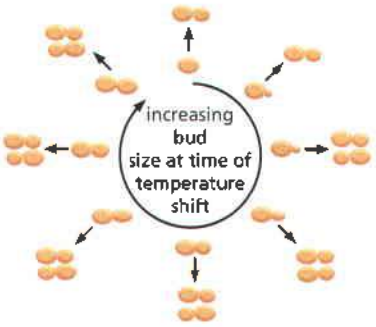


Figure Q17–1 Time-lapse photography of a temperature-sensitive mutant of yeast (Problem 17–9). Cells on the *inner ring* are arranged in order of their bud size, which corresponds to their position in the cell cycle. After 6 hours at 37°C, they have given rise to the cells shown on the *outer ring*. No further growth or division occurs.

expression at any point in the cell cycle. If expression is turned on at the beginning of S phase, all the cells divide satisfactorily and survive. By contrast, if Scc1 expression is turned on only after S phase is completed, the cells fail to divide and they die, even though Scc1 accumulates in the nucleus and interacts efficiently with chromosomes. Why do you suppose that cohesin must be present during S phase for cells to divide normally?

17–11 If cohesins join sister chromatids all along their length, how is it possible for condensins to generate mitotic chromosomes such as that shown in **Figure Q17–2**, which clearly shows the two sister chromatids as separate domains?

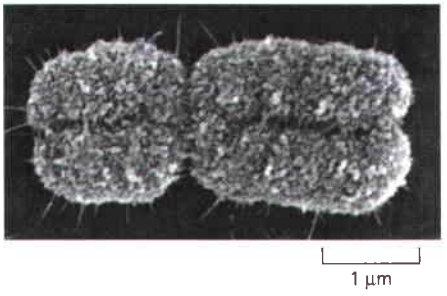


Figure Q17–2 A scanning electron micrograph of a fully condensed mitotic chromosome from vertebrate cells (Problem 17–11). (Courtesy of Terry D. Allen.)

17–12 High doses of caffeine interfere with the DNA replication checkpoint mechanism in mammalian cells. Why then do you suppose the Surgeon General has not yet issued an appropriate warning to heavy coffee and cola drinkers? A typical cup of coffee (150 mL) contains 100 mg of caffeine (196 g/mole). How many cups of coffee would you have to drink to reach the dose (10 mM) required to interfere with the DNA replication checkpoint mechanism? (A typical adult contains about 40 liters of water.)

17–13 A living cell from the lung epithelium of a newt is shown at different stages in M phase in **Figure Q17–3**. Order these light micrographs into the correct sequence and identify the stage in M phase that each represents.

17–14 How many kinetochores are there in a human cell at mitosis?

17–15 A classic paper clearly distinguished the properties of astral microtubules from those of kinetochore microtubules.

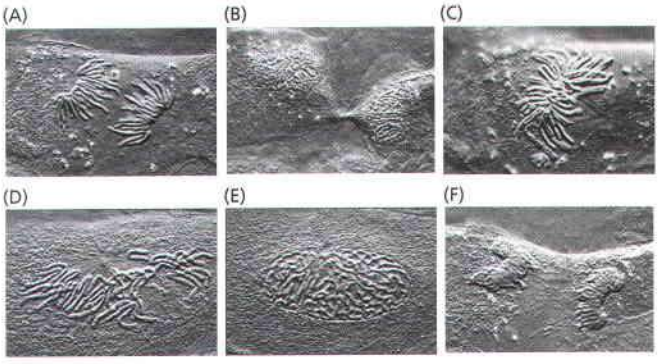


Figure Q17–3 Light micrographs of a single cell at different stages of M phase (Problem 17–13). (Courtesy of Conly L. Rieder.)

Centrosomes were used to initiate microtubule growth, and then chromosomes were added. The chromosomes bound to the free ends of the microtubules, as illustrated in **Figure Q17–4**. The complexes were then diluted to very low tubulin concentration (well below the critical concentration for microtubule assembly) and examined again (Figure Q17–4). As is evident, only the kinetochore microtubules were stable to dilution.

A. Why do you think kinetochore microtubules are stable?

B. How would you explain the disappearance of the astral microtubules after dilution? Do they detach from the centrosome, depolymerize from an end, or disintegrate along their length at random?

C. How would a time course after dilution help to distinguish among these possible mechanisms for disappearance of the astral microtubules?

17–16 What are the two distinct cytoskeletal machines that are assembled to carry out the mechanical processes of mitosis and cytokinesis in animal cells?

17–17 How do mitogens, growth factors, and survival factors differ from one another?

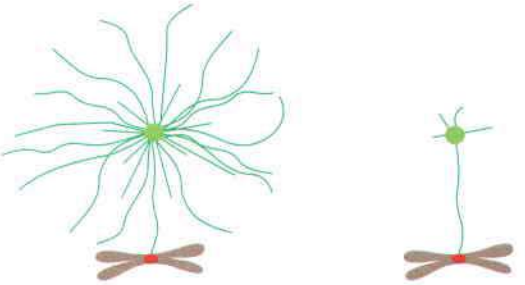


Figure Q17–4 Arrangements of centrosomes, chromosomes, and microtubules before and after dilution to low tubulin concentration (Problem 17–15).

REFERENCES

General

Morgan DO (2007) The Cell Cycle: Principles of Control. London: New Science Press.

Murray AW & Hunt T (1993) The Cell Cycle: An Introduction. New York: WH Freeman and Co.

Overview of the Cell Cycle

Forsburg SL & Nurse P (1991) Cell cycle regulation in the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Annu Rev Cell Biol* 7:227–256.

Hartwell LH, Culotti J, Pringle JR et al (1974) Genetic control of the cell division cycle in yeast. *Science* 183:46–51.

Kirschner M, Newport J & Gerhart J (1985) The timing of early developmental events in *Xenopus*. *Trends Genet* 1:41–47.

Nurse P, Thuriaux P & Nasmyth K (1976) Genetic control of the cell division cycle in the fission yeast *Schizosaccharomyces pombe*. *Mol Gen Genet* 146:167–178.

The Cell-Cycle Control System

Evans T, Rosenthal ET, Youngblom J et al (1983) Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. *Cell* 33:389–396.
Lohka MJ, Hayes MK & Maller JL (1988) Purification of maturation-promoting factor, an intracellular regulator of early mitotic events. *Proc Natl Acad Sci USA* 85:3009–3013.
Masui Y and Markert CL (1971) Cytoplasmic control of nuclear behavior during meiotic maturation of frog oocytes. *J Exp Zool* 177:129–146.
Morgan DO (1997) Cyclin-dependent kinases: engines, clocks, and microprocessors. *Annu Rev Cell Dev Biol* 13:261–291.
Murray AW & Kirschner MW (1989) Cyclin synthesis drives the early embryonic cell cycle. *Nature* 339:275–780.
Pavletich NP (1999) Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and CIP and Ink4 inhibitors. *J Mol Biol* 287:821–828.
Peters JM (2006) The anaphase promoting complex/cyclosome: a machine designed to destroy. *Nature Rev Mol Cell Biol* 7:644–656.
Petroski MD & Deshaies RJ (2005) Function and regulation of cullin-RING ubiquitin ligases. *Nature Rev Mol Cell Biol* 6:9–20.
Wittenberg C & Reed SI (2005) Cell cycle-dependent transcription in yeast: promoters, transcription factors, and transcriptomes. *Oncogene* 24:2746–2755.

S Phase

Arias EE & Walter JC (2007) Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev* 21:497–518.
Bell SP & Dutta A (2002) DNA replication in eukaryotic cells. *Annu Rev Biochem* 71:333–374.
Bell SP & Stillman B (1992) ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* 357:128–134.
Diffley JF (2004) Regulation of early events in chromosome replication. *Curr Biol* 14:R778–R786.
Groth A, Rocha W, Verreault A et al (2007) Chromatin challenges during DNA replication and repair. *Cell* 128:721–733.
Tanaka S, Umemori T, Hirai K et al (2007) CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature* 445:328–332.
Zegerman P & Diffley JF (2007) Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature* 445:281–285.

Mitosis

Cheeseman IM, Chappie JS, Wilson-Kubalek EM et al (2006) The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. *Cell* 127:983–997.
Dong Y, Vanden Beldt KJ, Meng X et al (2007) The outer plate in vertebrate kinetochores is a flexible network with multiple microtubule interactions. *Nature Cell Biol* 9:516–522.
Heald R, Tournebize R, Blank T et al (1996) Self-organization of microtubules into bipolar spindles around artificial chromosomes in *Xenopus* egg extracts. *Nature* 382:420–425.
Hirano T (2005) Condensins: organizing and segregating the genome. *Curr Biol* 15:R265–R275.
Kapoor TM, Lampson MA, Hergert P et al (2006) Chromosomes can congress to the metaphase plate before biorientation. *Science* 311:388–391.
Mitchison T & Kirschner M (1984) Dynamic instability of microtubule growth. *Nature* 312:237–242.
Mitchison TJ (1989) Polewards microtubule flux in the mitotic spindle: evidence from photoactivation of fluorescence. *J Cell Biol* 109:637–652.
Mitchison TJ & Salmon ED (2001) Mitosis: a history of division. *Nature Cell Biol* 3:E17–E21.
Musacchio A & Salmon ED (2007) The spindle-assembly checkpoint in space and time. *Nature Rev Mol Cell Biol* 8:379–393.
Nasmyth K (2002) Segregating sister genomes: the molecular biology of chromosome separation. *Science* 297:559–565.
Nigg EA (2007) Centrosome duplication: of rules and licenses. *Trends Cell Biol* 17:215–221.
Nurse P (1990) Universal control mechanism regulating onset of M-phase. *Nature* 344:503–508.
Page SL & Hawley RS (2003) Chromosome choreography: the meiotic ballet. *Science* 301:785–789.
Petronczki M, Siomos MF & Nasmyth K (2003) Un ménage à quatre: the molecular biology of chromosome segregation in meiosis. *Cell* 112:423–440.
Salmon ED (2005) Microtubules: a ring for the depolymerization motor. *Curr Biol* 15:R299–R302.
Tanaka TU, Stark MJ & Tanaka K (2005) Kinetochore capture and bi-orientation on the mitotic spindle. *Nature Rev Mol Cell Biol* 6:929–942.
Uhlmann F, Lottspeich F & Nasmyth K (1999) Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. *Nature* 400:37–42.
Wadsworth P & Khodjakov A (2004) E pluribus unum: towards a universal mechanism for spindle assembly. *Trends Cell Biol* 14:413–419.

Cytokinesis

Albertson R, Riggs B & Sullivan W (2005) Membrane traffic: a driving force in cytokinesis. *Trends Cell Biol* 15:92–101.
Burgess DR & Chang F (2005) Site selection for the cleavage furrow at cytokinesis. *Trends Cell Biol* 15:156–162.
Dechant R & Glotzer M (2003) Centrosome separation and central spindle assembly act in redundant pathways that regulate microtubule density and trigger cleavage furrow formation. *Dev Cell* 4:333–344.
Eggert US, Mitchison TJ & Field CM (2006) Animal cytokinesis: from parts list to mechanisms. *Annu Rev Biochem* 75:543–566.
Glotzer M (2005) The molecular requirements for cytokinesis. *Science* 307:1735–1739.
Grill SW, Gonczy P, Stelzer EH et al (2001) Polarity controls forces governing asymmetric spindle positioning in the *Caenorhabditis elegans* embryo. *Nature* 409:630–633.
Jurgens G (2005) Plant cytokinesis: fission by fusion. *Trends Cell Biol* 15:277–283.
Rappaport R (1986) Establishment of the mechanism of cytokinesis in animal cells. *Int Rev Cytol* 105:245–281.

Control of Cell Division and Cell Growth

Adhikary S & Eilers M (2005) Transcriptional regulation and transformation by Myc proteins. *Nature Rev Mol Cell Biol* 6:635–645.
Campisi J (2005) Senescent cells, tumor suppression, and organismal aging: good citizens, bad neighbors. *Cell* 120:513–522.
Conlon I & Raff M (1999) Size control in animal development. *Cell* 96:235–744.
Frolov MV, Huen DS, Stevaux O et al (2001) Functional antagonism between E2F family members. *Genes Dev* 15:2146–2160.
Harrison JC & Haber JE (2006) Surviving the breakup: the DNA damage checkpoint. *Annu Rev Genet* 40:209–235.
Jorgensen P & Tyers M (2004) How cells coordinate growth and division. *Curr Biol* 14:R1014–R1027.
Levine AJ (1997) p53, the cellular gatekeeper for growth and division. *Cell* 88:323–331.
Raff MC (1992) Social controls on cell survival and cell death. *Nature* 356:397–400.
Sherr CJ & DePinho RA (2000) Cellular senescence: mitotic clock or culture shock? *Cell* 102:407–410.
Sherr CJ & Roberts JM (1999) CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev* 13:1501–1512.
Trimarchi JM & Lees JA (2002) Sibling rivalry in the E2F family. *Nature Rev Mol Cell Biol* 3:11–20.
Vousden KH & Lu X (2002) Live or let die: the cell's response to p53. *Nature Rev Cancer* 2:594–604.
Zetterberg A & Larsson O (1985) Kinetic analysis of regulatory events in G1 leading to proliferation or quiescence of Swiss 3T3 cells. *Proc Natl Acad Sci USA* 82:5365–5369.

18 Apoptosis

Cell death plays a crucially important part in animal and plant development, and it usually continues into adulthood. In a healthy adult human, billions of cells die in the bone marrow and intestine every hour. Our tissues do not shrink because, by unknown regulatory mechanisms, cell division exactly balances the cell death. We now know that these "normal" cell deaths are suicides, in which the cells activate an intracellular death program and kill themselves in a controlled way—a process known as **programmed cell death**. The idea that animal cells have a built-in death program was proposed in the 1970s, but its general acceptance took another 20 years and depended on genetic studies in the nematode *C. elegans* that identified the first genes dedicated to programmed cell death and its control.

Programmed cell death in animals usually, but not exclusively, occurs by **apoptosis** (from the Greek word meaning "falling off," as leaves from a tree). Although apoptosis is only one form of programmed cell death, it is by far the most common and best understood, and, confusingly, biologists often use the terms programmed cell death and apoptosis interchangeably. Cells dying by apoptosis undergo characteristic morphological changes. <GCCC> They shrink and condense, the cytoskeleton collapses, the nuclear envelope disassembles, and the nuclear chromatin condenses and breaks up into fragments (**Figure 18–1**A). The cell surface often blebs and, if the cell is large, often breaks up into membrane-enclosed fragments called *apoptotic bodies*. Most importantly, the surface of the cell or apoptotic bodies becomes chemically altered, so that a neighboring cell or a macrophage (a specialized phagocytic cell, discussed in Chapter 23) rapidly engulfs them, before they can spill their contents (Figure 18–1B). In this way, the cell dies neatly and is rapidly cleared away, without causing a damaging inflammatory response. Because the cells are eaten and digested so quickly, there are usually few dead cells to be seen, even when large numbers of cells have died by apoptosis. This is probably why biologists overlooked apoptosis for many years and still probably underestimate its extent.

By contrast to apoptosis and other less well characterized forms of programmed cell death (which implies the operation of an intracellular death program), animal cells that die accidentally in response to an acute insult, such as trauma or a lack of blood supply, usually do so by a process called *cell necrosis*. Necrotic cells swell and burst, spilling their contents over their neighbors and eliciting an inflammatory response (Figure 18–1C).

Programmed cell death is not confined to animals. In plants, it occurs during development and in the senescence of flowers and leaves, as well as in the response to injury and infection. Programmed cell death even occurs in unicellular organisms, including yeasts and bacteria. The molecular mechanisms involved in these cases are distinct from those that mediate apoptosis in animal cells, and we shall not consider them. In this chapter, we discuss the functions of programmed cell death in animals, the molecular mechanism of apoptosis and its regulation, and how excessive or insufficient apoptosis can contribute to human disease.

Programmed Cell Death Eliminates Unwanted Cells

The amount of programmed cell death that occurs in developing and adult animal tissues can be astonishing. In the developing vertebrate nervous system, for

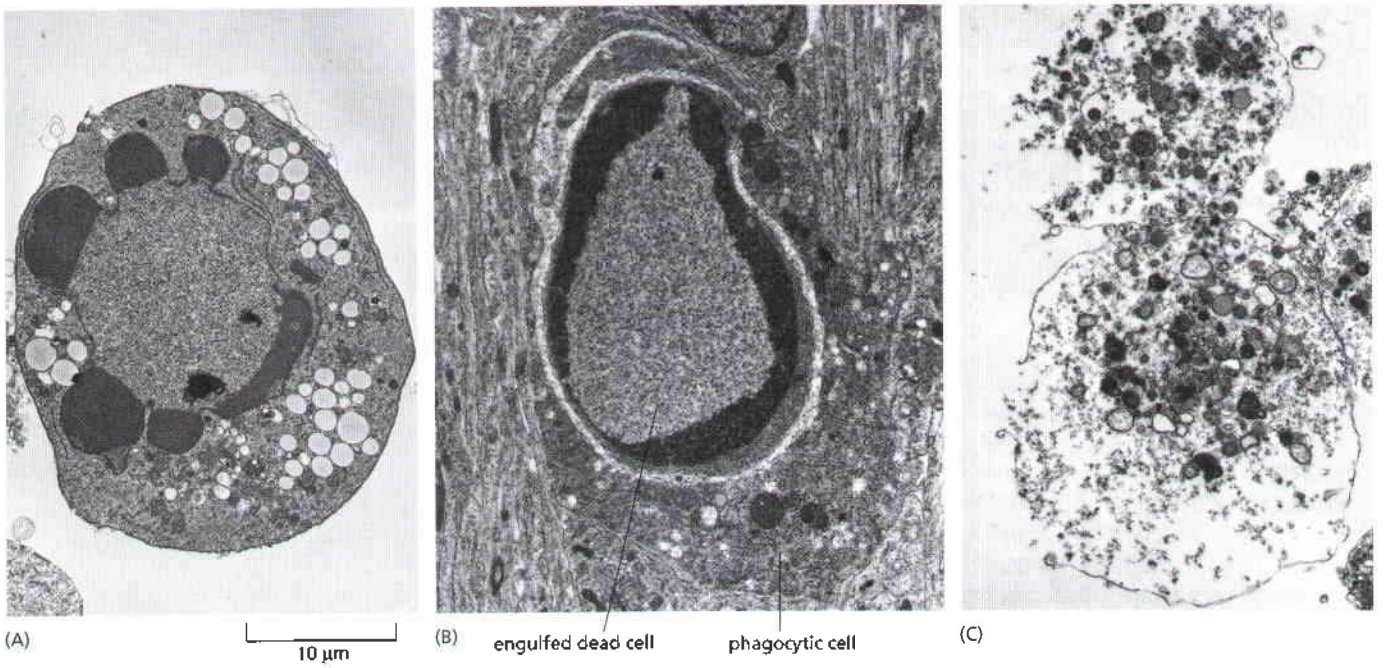


Figure 18–1 **Two distinct forms of cell death.** These electron micrographs show cells that have died by apoptosis (A and B) or by a type of accidental cell death called necrosis (C). The cells in (A) and (C) died in a culture dish, whereas the cell in (B) died in a developing tissue and has been engulfed by a phagocytic cell. Note that the cells in (A) and (B) have condensed but seem relatively intact, whereas the cell in (C) seems to have exploded. The large vacuoles visible in the cytoplasm of the cell in (A) are a variable feature of apoptosis. (Courtesy of Julia Burne.)

example, more than half of many types of nerve cells normally die soon after they are formed. It seems remarkably wasteful for so many cells to die, especially as the vast majority are perfectly healthy at the time they kill themselves. What purposes does this massive cell death serve?

In some cases, the answer is clear. In animal development, programmed cell death eliminates unwanted cells, usually by apoptosis. Cell death, for example, helps sculpt hands and feet during embryonic development: they start out as spade-like structures, and the individual digits separate only as the cells between them die, as illustrated for a mouse paw in Figure 18–2. In other cases, cells die when the structure they form is no longer needed. When a tadpole changes into a frog at metamorphosis, the cells in the tail die, and the tail, which is not needed in the frog, disappears (Figure 18–3). In many other cases, cell death helps regulate cell numbers. In the developing nervous system, for example, cell death adjusts the number of nerve cells to match the number of target cells that the nerve cells connect to, as we discuss later.

Programmed cell death also functions as a quality-control process in development, eliminating cells that are abnormal, misplaced, nonfunctional, or potentially dangerous to the animal. Striking examples occur in the vertebrate adaptive immune system, where apoptosis eliminates developing T and B lymphocytes that either fail to produce potentially useful antigen-specific receptors or produce self-reactive receptors that make the cells potentially dangerous; it also eliminates most of the lymphocytes activated by an infection, after they have helped destroy the responsible microbes (discussed in Chapter 25).

In adult tissues that are neither growing nor shrinking, cell death and cell division must be tightly regulated to ensure that they are exactly in balance. If

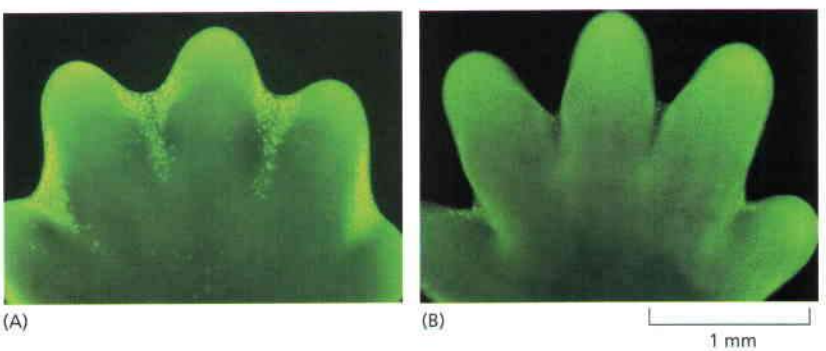


Figure 18–2 **Sculpting the digits in the developing mouse paw by apoptosis.** (A) The paw in this mouse fetus has been stained with a dye that specifically labels cells that have undergone apoptosis. The apoptotic cells appear as *bright green* dots between the developing digits. (B) The interdigital cell death has eliminated the tissue between the developing digits, as seen one day later, when there are very few apoptotic cells. (From W. Wood et al., *Development* 127:5245–5252, 2000. With permission from The Company of Biologists.)

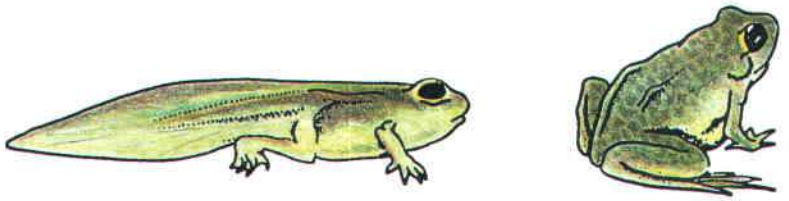


Figure 18–3 **Apoptosis during the metamorphosis of a tadpole into a frog.** As a tadpole changes into a frog, the cells in the tadpole tail are induced to undergo apoptosis; as a consequence, the tail is lost. An increase in thyroid hormone in the blood stimulates all the changes that occur during metamorphosis, including apoptosis in the tail.

part of the liver is removed in an adult rat, for example, liver cell proliferation increases to make up the loss. Conversely, if a rat is treated with the drug phenobarbital—which stimulates liver cell division (and thereby liver enlargement)—and then the phenobarbital treatment is stopped, apoptosis in the liver greatly increases until the liver has returned to its original size, usually within a week or so. Thus, the liver is kept at a constant size through the regulation of both the cell death rate and the cell birth rate, although the control mechanisms responsible for such regulation are largely unknown.

Apoptosis occurs at a staggeringly high rate in the adult human bone marrow, where most blood cells are produced. Here, for example, *neutrophils* (a type of white blood cell discussed in Chapter 23) are produced continuously in very large numbers, but the vast majority die by apoptosis in the bone marrow within a few days without ever functioning. This apparently futile cycle of production and destruction serves to maintain a ready supply of short-lived neutrophils that can be rapidly mobilized to fight infection wherever it occurs in the body. Compared with the life of the organism, cells are evidently cheap.

Animal cells can recognize damage in their various organelles and, if the damage is great enough, they can kill themselves by undergoing apoptosis. An important example is DNA damage, which can produce cancer-promoting mutations if not repaired. Cells have various ways of detecting DNA damage, and, if they cannot repair it, they often kill themselves by undergoing apoptosis.

Apoptotic Cells Are Biochemically Recognizable

Cells undergoing apoptosis not only have a characteristic morphology but also display characteristic biochemical changes, which can be used to identify apoptotic cells. During apoptosis, for example, an endonuclease cleaves the chromosomal DNA into fragments of distinctive sizes; because the cleavages occur in the linker regions between nucleosomes, the fragments separate into a characteristic ladder pattern when analyzed by gel electrophoresis (**Figure 18–4**A). Moreover, the cleavage of DNA generates many new DNA ends, which can be marked in apoptotic nuclei by using a labeled nucleotide in the so-called TUNEL technique (Figure 18–4B).

An especially important change occurs in the plasma membrane of apoptotic cells. The negatively charged phospholipid *phosphatidylserine* is normally exclusively located in the inner leaflet of the lipid bilayer of the plasma membrane (see Figures 10–3 and 10–16), but it flips to the outer leaflet in apoptotic cells, where it can serve as a marker of these cells. The phosphatidylserine on the surface of apoptotic cells can be visualized with a labeled form of the *Annexin V* protein, which specifically binds to this phospholipid. The cell-surface phosphatidylserine is more than a convenient marker of apoptosis for biologists; it helps signal to neighboring cells and macrophages to phagocytose the dying cell. In addition to serving as an "eat me" signal, it also blocks the inflammation often associated with phagocytosis: the phosphatidylserine-dependent engulfment of apoptotic cells inhibits the production of inflammation-inducing signal proteins (cytokines) by the phagocytic cell.

Macrophages will phagocytose most types of small particles, including oil droplets and glass beads, but they do not phagocytose any healthy cells in the animal, presumably because healthy cells express "don't eat me" signal molecules on their surface. Thus, in addition to expressing cell-surface "eat me" signals such as phosphatidylserine that stimulate phagocytosis, apoptotic cells must lose or inactivate their "don't eat me" signals in order for macrophages to ingest them.

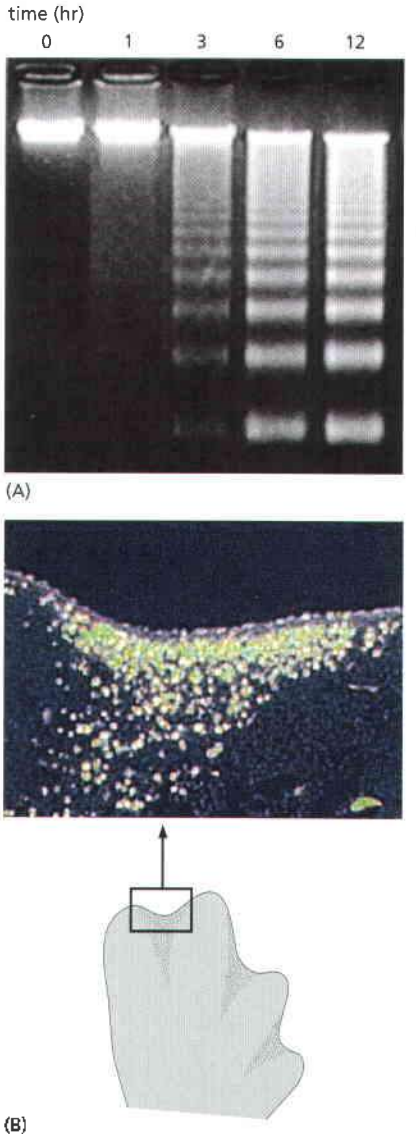


Figure 18–4 Markers of apoptosis. (A) Cleavage of nuclear DNA into a characteristic ladder pattern of fragments. Mouse thymus lymphocytes were treated with an antibody against the cell-surface death receptor Fas (discussed later), inducing the cells to undergo apoptosis. After various times (indicated in hours at the top of the figure), DNA was extracted, and the fragments were separated by size by electrophoresis in an agarose gel and stained with ethidium bromide. (B) The TUNEL technique was used to label the cut ends of DNA fragments in the nuclei of apoptotic cells in a tissue section of a developing chick leg bud; this cross section through the skin and underlying tissue is from a region between two developing digits, as indicated in the underlying drawing. The procedure is called the TUNEL (TdT-mediated dUTP nick end labeling) technique because the enzyme terminal deoxynucleotidyl transferase (TdT) adds chains of labeled deoxynucleotide (dUTP) to the 3′-OH ends of DNA fragments. (A, from D. McIlroy et al., *Genes Dev.* 14:549–558, 2000. With permisison from Cold Spring Harbor Laboratory Press; B, from V. Zuzarte-Luís and J.M. Hurlé, *Int. J. Dev. Biol.* 46:871–876, 2002. With permission from UBC Press.)

Cells undergoing apoptosis often lose the electrical potential that normally exists across the inner membrane of their mitochondria (discussed in Chapter 14). This membrane potential can be measured by the use of positively charged fluorescent dyes that accumulate in mitochondria, driven by the negative charge on the inside of the inner membrane. A decrease in the labeling of mitochondria with these dyes helps to identify cells that are undergoing apoptosis. As we discuss later, proteins such as *cytochrome c* are usually released from the space between the inner and outer membrane (the *intermembrane space*) of mitochondria during apoptosis, and the relocation of cytochrome *c* from mitochondria to the cytosol can be used as another marker of apoptosis (see Figure 18–7).

Apoptosis Depends on an Intracellular Proteolytic Cascade That Is Mediated by Caspases

The intracellular machinery responsible for apoptosis is similar in all animal cells. It depends on a family of proteases that have a cysteine at their active site and cleave their target proteins at specific aspartic acids. They are therefore called **caspases** (c for cysteine and asp for aspartic acid). Caspases are synthesized in the cell as inactive precursors, or **procaspases**, which are typically activated by proteolytic cleavage. Procaspase cleavage occurs at one or two specific aspartic acids and is catalyzed by other (already active) caspases; the procaspase is split into a large and a small subunit that form a heterodimer, and two such dimers assemble to form the active tetramer (Figure 18–5A). Once activated, caspases cleave, and thereby activate, other procaspases, resulting in an amplifying proteolytic cascade (Figure 18–5B).

Not all caspases mediate apoptosis. Indeed, the first caspase identified was a human protein called *interleukin-1-converting enzyme (ICE)*, which is concerned with inflammatory responses rather than with cell death; ICE cuts out the inflammation-inducing cytokine *interleukin-1 (IL1)* from a larger precursor protein. Subsequent to the discovery of ICE, a gene required for apoptosis in *C. elegans* was shown to encode a protein that is structurally and functionally similar to ICE, providing the first evidence that proteolysis and caspases are involved in apoptosis. It is now clear that, whereas several human caspases are involved in inflammatory and immune responses, most are involved in apoptosis (Table 18–1).

As shown in Figure 18–5B and Table 18–1, some of the procaspases that operate in apoptosis act at the start of the proteolytic cascade and are called **initiator procaspases**; when activated, they cleave and activate downstream **executioner procaspases**, which, then cleave and activate other executioner procaspases, as well as specific *target proteins* in the cell. Among the many target proteins cleaved by executioner caspases are the nuclear lamins (see Figure 18–5B), the cleavage of which causes the irreversible breakdown of the nuclear lamina (discussed in Chapter 16). Another target is a protein that normally holds

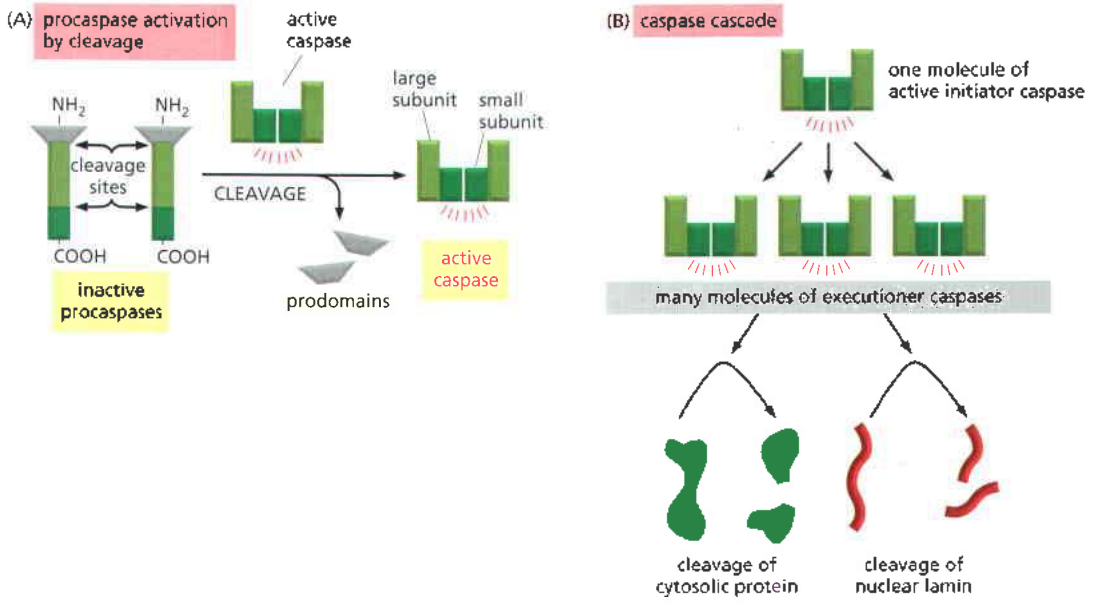


Figure 18–5 Procaspase activation during apoptosis. (A) Each caspase is initially made as an inactive proenzyme (procaspase). Some procaspases are activated by proteolytic cleavage by an activated caspase: two cleaved fragments from each of two procaspase molecules associate to form an active caspase, which is a tetramer of two small and two large subunits; the prodomains are usually discarded, as indicated. (B) The first procaspases activated are called *initiator procaspases*, which then cleave and activate many *executioner procaspase* molecules, producing an amplifying chain reaction (a proteolytic caspase cascade). The executioner caspases then cleave a variety of key proteins in the cell, including specific cytosolic proteins and nuclear lamins, as shown here, leading to the controlled death of the cell. Although not shown, the initiator procaspases are activated by adaptor proteins that bring the procaspases together in close proximity within an activation complex; although the initiator procaspases cleave each other within the complex, the cleavage serves only to stabilize the active protease.

the DNA-degrading enzyme mentioned earlier (an endonuclease) in an inactive form; its cleavage frees the endonuclease to cut up the DNA in the cell nucleus. Other target proteins include components of the cytoskeleton and cell–cell adhesion proteins that attach cells to their neighbors; the cleavage of these proteins helps the apoptotic cell to round up and detach from its neighbors, making it easier for a healthy neighboring cell to engulf it, or, in the case of an epithelial cell, for the neighbors to extrude the apoptotic cell from the cell sheet. The caspase cascade is not only destructive and self-amplifying but also irreversible, so that once a cell reaches a critical point along the path to destruction, it cannot turn back.

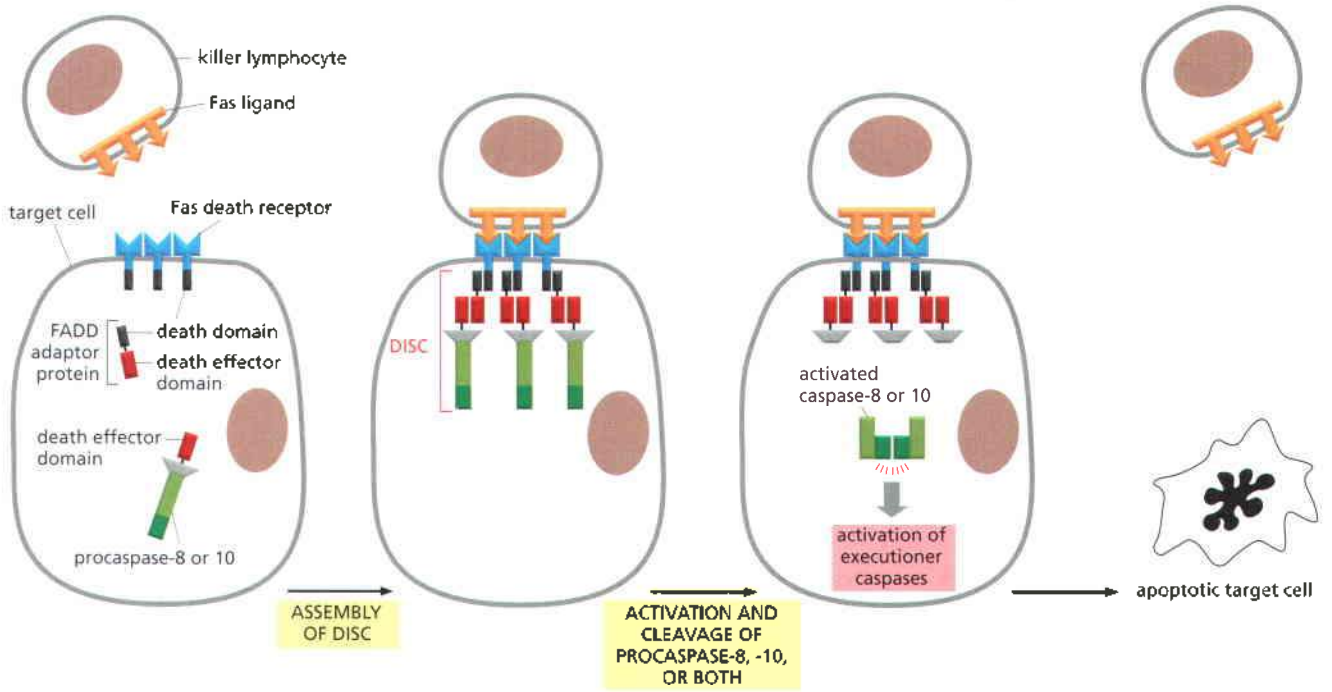
The caspases required for apoptosis vary depending on the cell type and stimulus. Inactivation of the mouse gene encoding caspase-3, an executioner caspase, for example, reduces normal apoptosis in the developing brain. As a result, the mouse often dies around birth with a deformed brain that contains too many cells. Apoptosis occurs normally, however, in many other organs of such mice.

From the earliest stages of an animal's development, healthy cells continuously make the procaspases and other proteins required for apoptosis. Thus, the apoptosis machinery is always in place; all that is needed is a trigger to activate it. How, then, is a caspase cascade initiated? In particular, how is the first procaspase in the cascade activated? Initiator procaspases have a long *prodomain*, which contains a *caspase recruitment domain (CARD)* that enables them to assemble with adaptor proteins into *activation complexes* when the cell receives a signal to undergo apoptosis. Once incorporated into such a complex, the initiator procaspases are brought into close proximity, which is sufficient to activate them; they then cleave each other to make the process irreversible. The activated initiator caspases then cleave and activate executioner procaspases, thereby initiating a proteolytic caspase cascade, which amplifies the death signal and spreads it throughout the cell.

The two best understood signaling pathways that can activate a caspase cascade leading to apoptosis in mammalian cells are called the *extrinsic pathway* and the *intrinsic pathway*. Each uses its own initiator procaspases and activation complex, as we now discuss.

Table 18–1 Some Human Caspases

Caspases involved in inflammation	caspases 1 (ICE), 4, 5
Caspases involved in apoptosis	
Initiator caspases	caspases 2, 8, 9, 10
Executioner caspases	caspases 3, 6, 7



Cell-Surface Death Receptors Activate the Extrinsic Pathway of Apoptosis

Extracellular signal proteins binding to cell-surface **death receptors** trigger the **extrinsic pathway** of apoptosis. Death receptors are transmembrane proteins containing an extracellular ligand-binding domain, a single transmembrane domain, and an intracellular *death domain*, which is required for the receptors to activate the apoptotic program. The receptors are homotrimers and belong to the *tumor necrosis factor (TNF) receptor* family, which includes a receptor for TNF itself (discussed in Chapter 15) and the *Fas* death receptor. The ligands that activate the death receptors are also homotrimers; they are structurally related to one another and belong to the *TNF family* of signal proteins.

A well-understood example of how death receptors trigger the extrinsic pathway of apoptosis is the activation of **Fas** on the surface of a target cell by **Fas ligand** on the surface of a killer (cytotoxic) lymphocyte (discussed in Chapter 25). When activated by the binding of Fas ligand, the death domains on the cytosolic tails of the Fas death receptors recruit intracellular adaptor proteins, which in turn recruit initiator procaspases (*procaspase-8*, *procaspase-10*, or both), forming a **death-inducing signaling complex (DISC)**. Once activated in the DISC, the initiator caspases activate downstream executioner procaspases to induce apoptosis (**Figure 18–6**). As we discuss later, in some cells the extrinsic pathway must recruit the intrinsic apoptotic pathway to amplify the caspase cascade in order to kill the cell.

Many cells produce inhibitory proteins that act either extracellularly or intracellularly to restrain the extrinsic pathway. For example, some produce cell-surface *decoy receptors*, which have a ligand-binding domain but not a death domain; because they can bind a death ligand but cannot activate apoptosis, the decoys competitively inhibit the death receptors. Cells can also produce intracellular blocking proteins such as *FLIP*, which resembles an initiator procaspase but lacks the proteolytic domain; it competes with procaspase-8 and procaspase-10 for binding sites in the DISC and thereby inhibits the activation of these initiator procaspases. Such inhibitory mechanisms help prevent the inappropriate activation of the extrinsic pathway of apoptosis.

In some circumstances, death receptors activate other intracellular signaling pathways that do not lead to apoptosis. TNF receptors, for example, can also activate the NFκB pathway (discussed in Chapter 15), which can promote cell

Figure 18–6 The extrinsic pathway of apoptosis activated through Fas death receptors. Fas ligand on the surface of a killer lymphocyte activates Fas death receptors on the surface of the target cell. Both the ligand and receptor are homotrimers. The cytosolic tail of Fas then recruits the adaptor protein FADD via the death domain on each protein (FADD stands for Fas-associated death domain). Each FADD protein then recruits an initiator procaspase (procaspase-8, procaspase-10, or both) via a death effector domain on both FADD and the procaspase, forming a death-inducing signaling complex (DISC). Within the DISC, the initiator procaspase molecules are brought into close proximity, which activates them; the activated procaspases then cleave one another to stabilize the activated protease, which is now a caspase. Activated caspase-8 and caspase-10 then cleave and activate executioner procaspases, producing a caspase cascade, which leads to apoptosis.

survival and activate genes involved in inflammatory responses. Which responses dominate depends on the type of cell and the other signals acting on it.

The Intrinsic Pathway of Apoptosis Depends on Mitochondria

Cells can also activate their apoptosis program from inside the cell, usually in response to injury or other stresses, such as DNA damage or lack of oxygen, nutrients, or extracellular survival signals (discussed later). In vertebrate cells, such intracellular activation of the apoptotic death program occurs via the **intrinsic pathway** of apoptosis, which depends on the release into the cytosol of mitochondrial proteins that normally reside in the intermembrane space of these organelles (see Figure 12–21A). Some of the released proteins activate a caspase proteolytic cascade in the cytoplasm, leading to apoptosis.

A crucial protein released from mitochondria in the intrinsic pathway is **cytochrome *c***, a water-soluble component of the mitochondrial electron-transport chain. When released into the cytosol (**Figure 18–7**), it has an entirely different function: it binds to a procaspase-activating adaptor protein called **Apaf1** *(apoptotic protease activating factor-1)*, causing the Apaf1 to oligomerize into a wheel-like heptamer called an **apoptosome**. The Apaf1 proteins in the apoptosome then recruit initiator procaspase proteins *(procaspase-9)*, which are activated by proximity in the apoptosome, just as procaspase-8 and -10 proteins are activated in the DISC. The activated caspase-9 molecules then activate downstream executioner procaspases to induce apoptosis (**Figure 18–8**).

As mentioned earlier, in some cells, the extrinsic pathway must recruit the intrinsic pathway to amplify the apoptotic signal to kill the cell. It does so by activating a member of the *Bcl2* family of proteins, which we now discuss.

Bcl2 Proteins Regulate the Intrinsic Pathway of Apoptosis

The intrinsic pathway of apoptosis is tightly regulated to ensure that cells kill themselves only when it is appropriate. A major class of intracellular regulators of apoptosis is the **Bcl2** family of proteins, which, like the caspase family, has been conserved in evolution from worms to humans; a human Bcl2 protein, for example, can suppress apoptosis when expressed in *C. elegans*.

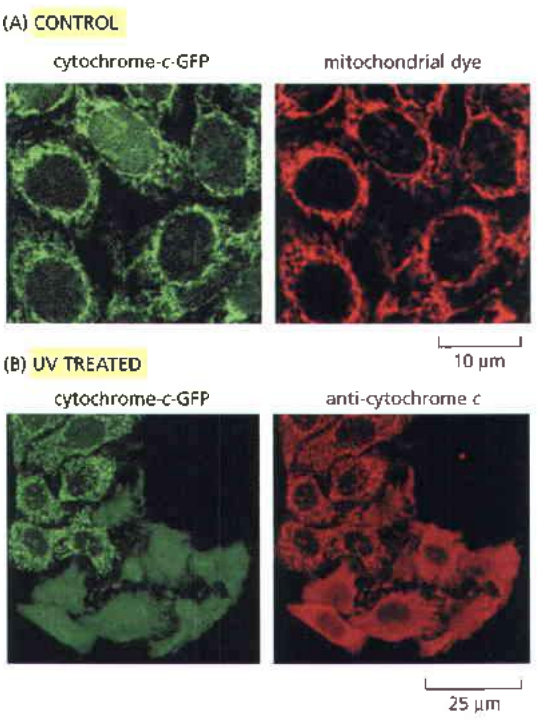
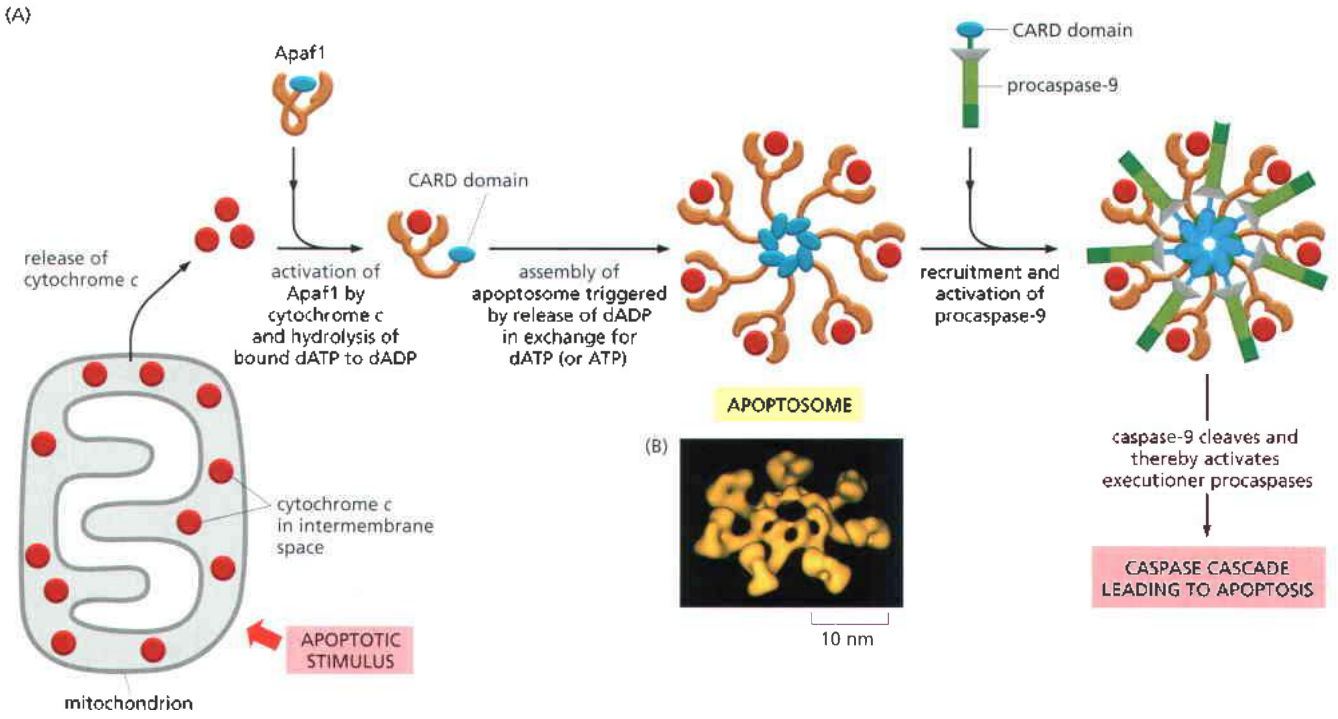


Figure 18–7 Release of cytochrome *c* from mitochondria during apoptosis. Fluorescence micrographs of human cancer cells in culture. (A) The control cells were transfected with a gene encoding a fusion protein consisting of cytochrome *c* linked to green fluorescent protein (cytochrome-*c*-GFP); they were also treated with a positively charged red dye that accumulates in mitochondria. The overlapping distribution of the *green* and *red* indicate that the cytochrome-*c*-GFP is located in mitochondria. (B) Cells expressing cytochrome-*c*-GFP were irradiated with ultraviolet light to induce apoptosis, and after 5 hours they were stained with antibodies (in *red*) against cytochrome *c*; the cytochrome-*c*-GFP is also shown (in *green*). The six cells in the bottom half of the micrographs in B have released their cytochrome *c* from mitochondria into the cytosol, whereas the cells in the upper half of the micrographs have not yet done so. (From J.C. Goldstein et al., *Nat. Cell Biol.* 2:156–162, 2000. With permission from Macmillan Publishers Ltd.)



Mammalian Bcl2 proteins regulate the intrinsic pathway of apoptosis mainly by controlling the release of cytochrome *c* and other intermembrane mitochondrial proteins into the cytosol. Some Bcl2 proteins are *pro-apoptotic* and promote apoptosis by enhancing the release, whereas others are *anti-apoptotic* and inhibit apoptosis by blocking the release. The pro-apoptotic and anti-apoptotic Bcl2 proteins can bind to each other in various combinations to form heterodimers, in which the two proteins inhibit each other's function. The balance between the activities of these two functional classes of Bcl2 proteins largely determines whether a mammalian cell lives or dies by the intrinsic pathway of apoptosis.

As illustrated in **Figure 18–9**, the anti-apoptotic Bcl2 proteins, including *Bcl2* itself (the founding member of the Bcl2 family) and *Bcl-X_L*, share four distinctive *Bcl2 homology (BH) domains* (BH1–4). The pro-apoptotic Bcl2 proteins consist of two subfamilies—the *BH123* proteins and the *BH3-only* proteins. The main BH123 proteins are *Bax* and *Bak*, which are structurally similar to Bcl2 but lack the BH4 domain. The BH3-only proteins share sequence homology with Bcl2 in only the BH3 domain (see Figure 18–9).

When an apoptotic stimulus triggers the intrinsic pathway, the pro-apoptotic **BH123 proteins** become activated and aggregate to form oligomers in the mitochondrial outer membrane, inducing the release of cytochrome *c* and other intermembrane proteins by an unknown mechanism (**Figure 18–10**). In mammalian cells, **Bax** and **Bak** are the main BH123 proteins, and at least one of them is required for the intrinsic pathway of apoptosis to operate: mutant mouse cells that lack both proteins are resistant to all pro-apoptotic signals that normally activate this pathway. Whereas Bak is tightly bound to the mitochondrial outer membrane even in the absence of an apoptotic signal, Bax is mainly located in the cytosol and translocates to the mitochondria only after an apoptotic signal activates it. As we discuss below, the activation of Bax and Bak usually depends on activated pro-apoptotic BH3-only proteins. Both Bax and Bak also operate on the surface of the endoplasmic reticulum (ER) and nuclear membranes; when activated in response to ER stress, they are thought to release Ca^{2+} into the cytosol, which helps activate the mitochondrial-dependent intrinsic pathway of apoptosis by a poorly understood mechanism.

The **anti-apoptotic Bcl2 proteins** such as **Bcl2** itself and **Bcl-X_L** are also mainly located on the cytosolic surface of the outer mitochondrial membrane, the ER, and the nuclear envelope, where they help preserve the integrity of the

Figure 18–8 The intrinsic pathway of apoptosis. (A) A schematic drawing of how cytochrome *c* released from mitochondria activates Apaf1. The binding of cytochrome *c* causes the Apaf1 to hydrolyze its bound dATP to dADP (not shown). The replacement of the dADP with dATP or ATP (not shown) then induces the complex of Apaf1 and cytochrome *c* to aggregate to form a large, heptameric apoptosome, which then recruits procaspase-9 through a caspase recruitment domain (CARD) in each protein. The procaspase-9 molecules are activated within the apoptosome and are now able to cleave and activate downstream executioner procaspases, which leads to the cleavage and activation of these molecules in a caspase cascade. Other proteins released from the mitochondrial intermembrane space are not shown. (B) A model of the three-dimensional structure of an apoptosome. Note that some scientists use the term "apoptosome" to refer to the complex containing procaspase-9. (B, from D. Aceham et al., *Mol. Cell* 9:423–432, 2002. With permission from Elsevier.)

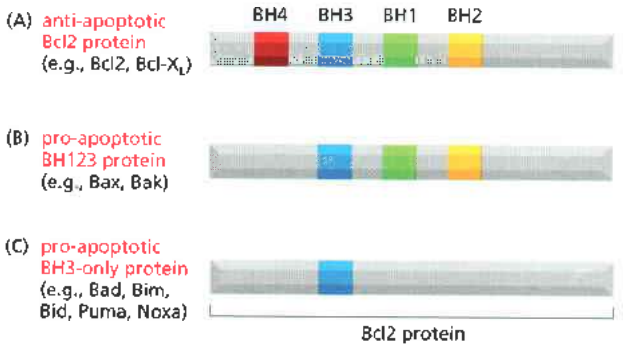


Figure 18–9 **The three classes of Bcl2 proteins.** Note that the BH3 domain is the only BH domain shared by all Bcl2 family members; it mediates the direct interactions between pro-apoptotic and anti-apoptotic family members.

membrane—preventing, for example, inappropriate release of intermembrane proteins from mitochondria and of Ca^{2+} from the ER. These proteins inhibit apoptosis mainly by binding to and inhibiting pro-apoptotic Bcl2 proteins—either on these membranes or in the cytosol. On the outer mitochondrial membrane, for example, they bind to Bak and prevent it from oligomerizing, thereby inhibiting the release of cytochrome *c* and other intermembrane proteins. There are at least five mammalian anti-apoptotic Bcl2 proteins, and every mammalian cell requires at least one to survive. Moreover, a number of these proteins must be inhibited for the intrinsic pathway to induce apoptosis; the BH3-only proteins mediate the inhibition.

The **BH3-only proteins** are the largest subclass of Bcl2 family proteins. The cell either produces or activates them in response to an apoptotic stimulus, and they are thought to promote apoptosis mainly by inhibiting anti-apoptotic Bcl2 proteins. Their BH3 domain binds to a long hydrophobic groove on anti-apoptotic Bcl2 proteins, neutralizing their activity. By a poorly understood mechanism, this binding and inhibition enables the aggregation of Bax and Bak on the surface of mitochondria, which triggers the release of the intermembrane mitochondrial proteins that induce apoptosis (**Figure 18–11**). Some BH3-only proteins may bind directly to Bax and Bak to help trigger the activation and aggregation of these BH123 pro-apoptotic proteins on mitochondria and thereby help release the intermembrane proteins.

BH3-only proteins provide the crucial link between apoptotic stimuli and the intrinsic pathway of apoptosis, with different stimuli activating different BH3-only proteins. When some cells are deprived of extracellular survival signals, for example, an intracellular signaling pathway that depends on the MAP kinase *JNK* activates the transcription of the gene encoding the BH3-only protein *Bim*, which then triggers the intrinsic pathway. Similarly, in response to DNA damage that cannot be repaired, the tumor suppressor protein **p53** accumulates (discussed in Chapters 17 and 20) and activates the transcription of genes that encode the BH3-only proteins *Puma* and *Noxa*; these BH3-only proteins then trigger the intrinsic pathway, thereby eliminating a potentially dangerous cell that could otherwise become cancerous.

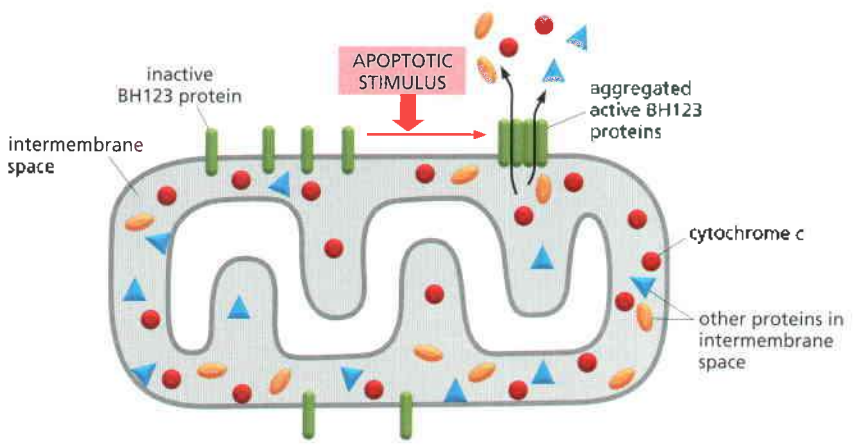


Figure 18–10 **The role of BH123 pro-apoptotic Bcl2 proteins (mainly Bax and Bak) in the release of mitochondrial intermembrane proteins in the intrinsic pathway of apoptosis.** When activated by an apoptotic stimulus, the BH123 proteins aggregate on the outer mitochondrial membrane and release cytochrome *c* and other proteins from the intermembrane space into the cytosol by an unknown mechanism.

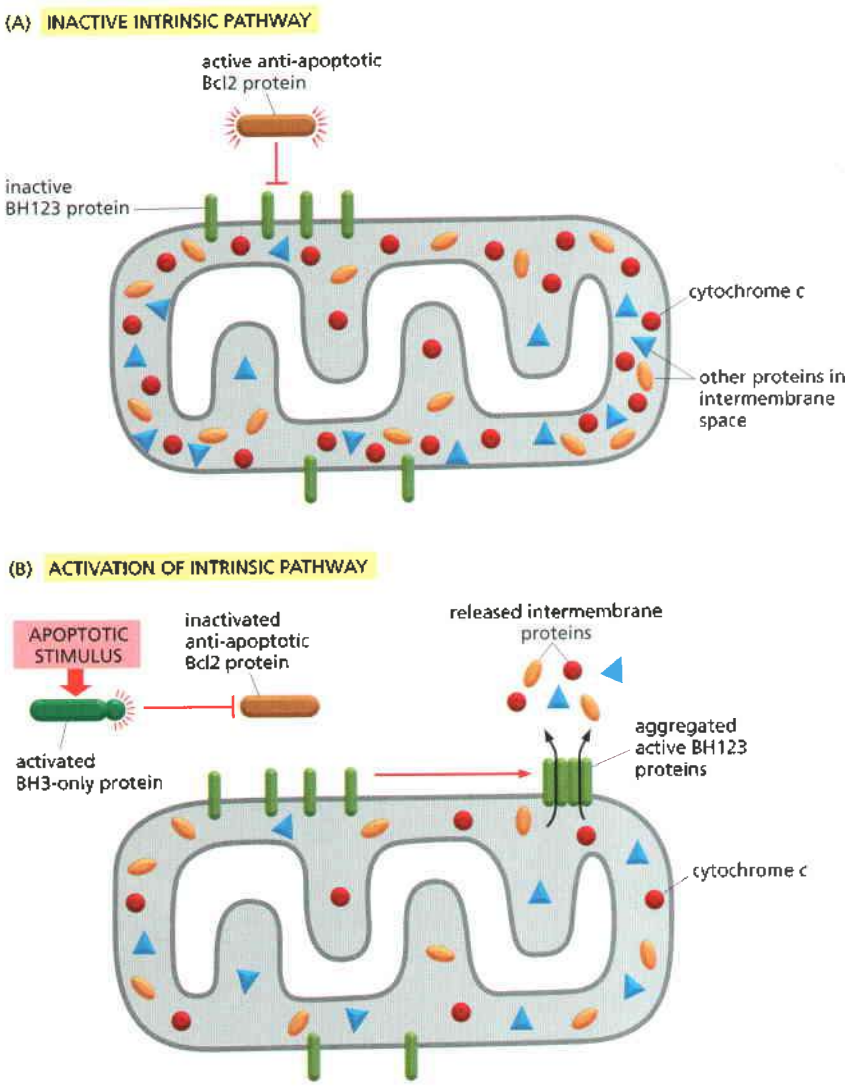


Figure 18–11 **How pro-apoptotic BH3-only and anti-apoptotic Bcl2 proteins regulate the intrinsic pathway of apoptosis.** (A) In the absence of an apoptotic stimulus, anti-apoptotic Bcl2 proteins bind to and inhibit the BH123 proteins on the mitochondrial outer membrane (and in the cytosol—not shown). (B) In the presence of an apoptotic stimulus, BH3-only proteins are activated and bind to the anti-apoptotic Bcl2 proteins so that they can no longer inhibit the BH123 proteins, which now become activated and aggregate in the outer mitochondrial membrane and promote the release of intermembrane mitochondrial proteins into the cytosol. Some activated BH3-only proteins may stimulate mitochondrial protein release more directly by binding to and activating the BH123 proteins. Although not shown, the anti-apoptotic Bcl2 proteins are bound to the mitochondrial surface.

As mentioned earlier, in some cells the extrinsic apoptotic pathway recruits the intrinsic pathway to amplify the caspase cascade to kill the cell. The BH3-only protein *Bid* is the link between the two pathways. When death receptors activate the extrinsic pathway in these cells, the initiator caspase, caspase-8, cleaves Bid, producing a truncated form of Bid called *tBid*. tBid translocates to mitochondria, where it inhibits anti-apoptotic Bcl2 proteins and triggers the aggregation of pro-apoptotic BH123 proteins to release cytochrome *c* and other intermembrane proteins, thereby amplifying the death signal.

The BH3-only proteins Bid, Bim, and Puma (see Figure 18–9) can inhibit all of the anti-apoptotic Bcl2 proteins, whereas the other BH3-only proteins can inhibit only a small subset of the anti-apoptotic proteins. Thus, Bid, Bim, and Puma are the most potent activators of apoptosis in the BH3-only subfamily of Bcl2 proteins.

Bcl2 proteins are not the only intracellular regulators of apoptosis. The IAP (inhibitor of apoptosis) proteins also play an important part in suppressing apoptosis, especially in *Drosophila*.

IAPs Inhibit Caspases

Inhibitors of apoptosis (IAPs) were first identified in certain insect viruses (baculoviruses), which encode IAP proteins to prevent a host cell that is infected by the virus from killing itself by apoptosis. (Virus-infected animal cells frequently kill themselves to prevent the virus from replicating and infecting other cells.) It is now known that most animal cells also make IAP proteins.

All IAPs have one or more BIR (baculovirus IAP repeat) domains, which enable them to bind to and inhibit activated caspases. Some IAPs also polyubiquitylate caspases, marking the caspases for destruction by proteasomes. In this way, the IAPs set an inhibitory threshold that activated caspases must overcome to trigger apoptosis.

In *Drosophila* at least, this inhibitory barrier provided by IAPs can be neutralized by **anti-IAP** proteins, which are produced in response to various apoptotic stimuli. There are five anti-IAPs in flies, including *Reaper*, *Grim*, and *Hid*, and their only structural similarity is their short, N-terminal, IAP-binding motif, which binds to the BIR domain of IAPs, preventing the domain from binding to a caspase. Deletion of the three genes encoding Reaper, Grim, and Hid blocks apoptosis in flies. Conversely, inactivation of one of the two genes that encode IAPs in *Drosophila* causes all of the cells in the developing fly embryo to undergo apoptosis. Clearly, the balance between IAPs and anti-IAPs is tightly regulated and is crucial for controlling apoptosis in the fly.

The role of mammalian anti-IAP proteins in apoptosis is more controversial. As illustrated in **Figure 18–12**, anti-IAPs are released from the mitochondrial intermembrane space when the intrinsic pathway of apoptosis is activated, blocking IAPs in the cytosol and thereby promoting apoptosis. When, however, the genes encoding two known mammalian anti-IAPs called *Smac* (also called *DIABLO*) and *Omi* are inactivated in mouse cells, apoptosis is apparently unaffected—hence the controversy about their normal roles in regulating apoptosis.

In summary, the combined activities of the Bcl2 proteins, IAPs, and anti-IAPs determine the sensitivity of an animal cell to apoptosis-inducing stimuli, with IAPs and anti-IAPs dominant in flies and Bcl2 proteins dominant in mammals.

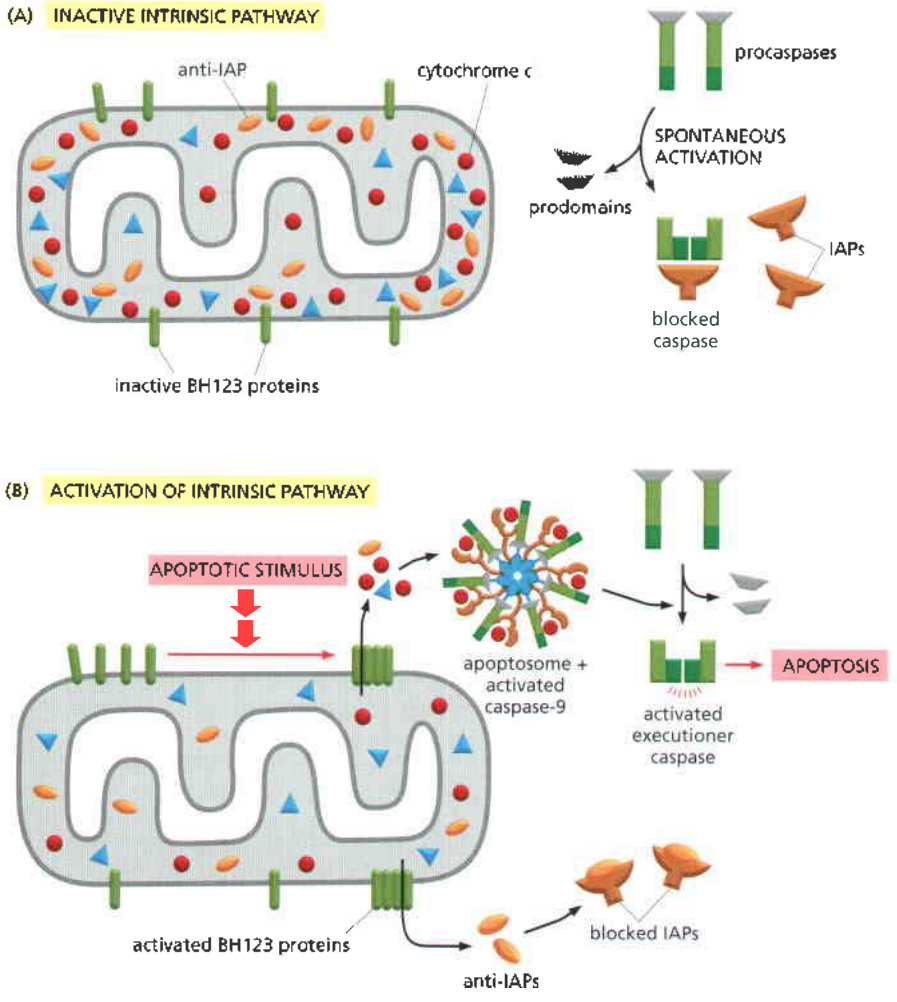


Figure 18–12 A proposed model for the roles of IAPs and anti-IAPs in the control of apoptosis in mammalian cells. (A) In the absence of an apoptotic stimulus, IAPs prevent accidental apoptosis caused by the spontaneous activation of procaspases. The IAPs are located in the cytosol and bind to and inhibit any spontaneously activated caspases. Some IAPs are also ubiquitin ligases that ubiquitylate the caspases they bind to, marking them for degradation in proteasomes (not shown). (B) When an apoptotic stimulus activates the intrinsic pathway, among the proteins released from the mitochondrial intermembrane space are anti-IAP proteins, which bind to and block the inhibitory activity of the IAPs. At the same time, the released cytochrome *c* triggers the assembly of apoptosomes, which can now activate a caspase cascade, leading to apoptosis.

Extracellular Survival Factors Inhibit Apoptosis in Various Ways

As discussed in Chapter 15, intercellular signals regulate most activities of animal cells, including apoptosis. These extracellular signals are part of the normal "social" controls that ensure that individual cells behave for the good of the organism as a whole—in this case, by surviving when they are needed and killing themselves when they are not. Some extracellular signal molecules stimulate apoptosis, whereas others inhibit it. We have discussed signal proteins such as Fas ligand that activate death receptors and thereby trigger the extrinsic pathway of apoptosis. Other extracellular signal molecules that stimulate apoptosis are especially important during animal development: a surge of thyroid hormone in the bloodstream, for example, signals cells in the tadpole tail to undergo apoptosis at metamorphosis (see Figure 18–3), while locally produced bone morphogenic proteins (BMPs, discussed in Chapters 15 and 22) stimulate cells between developing fingers and toes to kill themselves (see Figure 18–2). Here, however, we focus on extracellular signal molecules that inhibit apoptosis, which are collectively called **survival factors**.

Most animal cells require continuous signaling from other cells to avoid apoptosis. This surprising arrangement apparently helps ensure that cells survive only when and where they are needed. Nerve cells, for example, are produced in excess in the developing nervous system and then compete for limited amounts of survival factors that are secreted by the target cells that they normally connect to. Nerve cells that receive enough of the appropriate type of survival signal live, while the others die. In this way, the number of surviving neurons is automatically adjusted so that it is appropriate for the number of target cells they connect with (**Figure 18–13**). A similar competition for limited amounts of survival factors produced by neighboring cells is thought to control cell numbers in other tissues, both during development and in adulthood.

Survival factors usually bind to cell-surface receptors, which activate intracellular signaling pathways that suppress the apoptotic program, often by regulating members of the Bcl2 family of proteins. Some survival factors, for example, stimulate an increased production of anti-apoptotic Bcl2 proteins such as Bcl2 itself or Bcl-X_L (**Figure 18–14**A). Others act by inhibiting the function of BH3-only pro-apoptotic Bcl2 proteins such as *Bad* (Figure 18–14B). In *Drosophila*, some survival factors act by phosphorylating and inactivating anti-IAP proteins, thereby enabling IAP proteins to suppress apoptosis (Figure 18–14C).

When mammalian cells are deprived of survival factors. they kill themselves by producing and activating pro-apoptotic BH3-only proteins, which activate the intrinsic pathway of apoptosis by overriding the anti-apoptotic Bcl2 proteins that are required to keep the cells alive. Mouse cells that lack both Bax and Bak are unable to activate the intrinsic pathway and can therefore live for weeks in culture in the absence of survival factors; without survival signals, however, the cells cannot efficiently import nutrients. Such cells fuel their metabolic needs through *autophagy*, in which the cell sequesters organelles and bits of its cytoplasm within autophagosomes, which then fuse with lysosomes (discussed in Chapter 13). The cells eventually die from starvation, but not by apoptosis.

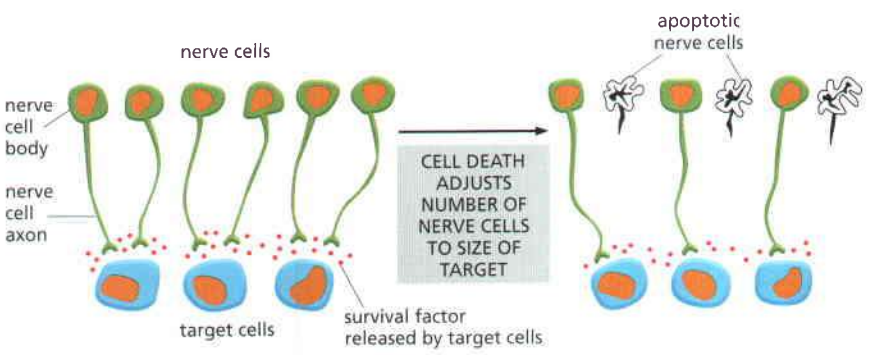


Figure 18–13 The role of survival factors and cell death in adjusting the number of developing nerve cells to the amount of target tissue. More nerve cells are produced than can be supported by the limited amount of survival factors released by the target cells. Therefore, some cells receive an insufficient amount of survival factors to avoid apoptosis. This strategy of overproduction followed by culling ensures that all target cells are contacted by nerve cells and that the extra nerve cells are automatically eliminated.

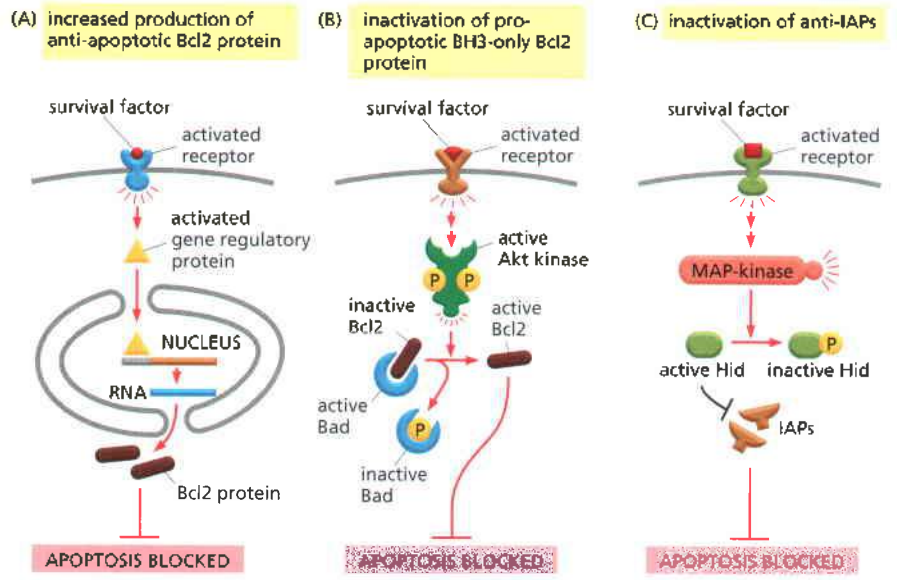


Figure 18–14 Three ways that extracellular survival factors can inhibit apoptosis. (A) Some survival factors suppress apoptosis by stimulating the transcription of genes that encode anti-apoptotic Bcl2 proteins such as Bcl2 itself or Bcl-X_L. (B) Many others activate the serine/threonine protein kinase Akt, which, among many other targets, phosphorylates and inactivates the BH3-only pro-apoptotic Bcl2 protein Bad (see Figure 15–64). When not phosphorylated, Bad promotes apoptosis by binding to and inhibiting Bcl2; once phosphorylated, Bad dissociates, freeing Bcl2 to suppress apoptosis. Akt also suppresses apoptosis by phosphorylating and inactivating gene regulatory proteins of the Forkhead family that stimulate the transcription of genes encoding proteins that promote apoptosis (not shown). (C) In *Drosophila*, some survival factors inhibit apoptosis by stimulating the phosphorylation of the anti-IAP protein Hid. When not phosphorylated, Hid promotes cell death by inhibiting IAPs. Once phosphorylated, Hid no longer inhibits IAPs, which become active and block apoptosis.

Either Excessive or Insufficient Apoptosis Can Contribute to Disease

There are many human disorders in which excessive numbers of cells undergo apoptosis and thereby contribute to tissue damage. Among the most dramatic examples are heart attacks and strokes. In these acute conditions, many cells die by necrosis as a result of ischemia (inadequate blood supply), but some of the less affected cells die by apoptosis. It is hoped that, in the future, drugs such as caspase inhibitors that block apoptosis will prove useful in saving cells in these conditions.

There are other conditions where too few cells die by apoptosis. Mutations in mice and humans, for example, that inactivate the genes that encode the Fas death receptor or the Fas ligand prevent the normal death of some lymphocytes, causing these cells to accumulate in excessive numbers in the spleen and lymph glands. In many cases, this leads to autoimmune disease, in which the lymphocytes react against the individual's own tissues.

Decreased apoptosis also makes an important contribution to many tumors, as cancer cells often regulate the apoptotic program abnormally. <TGAA> The *Bcl2* gene, for example, was first identified in a common form of lymphocyte cancer in humans, where a chromosome translocation causes excessive production of the Bcl2 protein; indeed, Bcl2 gets its name from this *B cell lymphoma*. The high level of Bcl2 protein in the lymphocytes that carry the translocation promotes the development of cancer by inhibiting apoptosis, thereby prolonging cell survival and increasing cell numbers; it also decreases the cells' sensitivity to anticancer drugs, which commonly work by causing cancer cells to undergo apoptosis.

Similarly, the gene encoding the tumor suppressor protein p53 is mutated in 50% of human cancers so that it no longer promotes apoptosis or cell-cycle arrest in response to DNA damage. The lack of p53 function therefore enables the cancer cells to survive and proliferate even when their DNA is damaged; in this way, the cells accumulate more mutations, some of which make the cancer more malignant (discussed in Chapter 20). As many anticancer drugs induce apoptosis (and cell-cycle arrest) by a p53-dependent mechanism (discussed in Chapters 17 and 20), the loss of p53 function also makes cancer cells less sensitive to these drugs.

Most human cancers arise in epithelial tissues such as those in the lung, intestinal tract, breast, and prostate. Such cancer cells display many abnormalities in their behavior, including a decreased ability to adhere to the extracellular matrix and to adhere to one another at specialized cell–cell junctions. In the next chapter, we discuss the remarkable structures and functions of the extracellular matrix and cell junctions.

Summary

Cells can activate an intracellular death program and kill themselves in a controlled way—a process called programmed cell death. In this way, animal cells that are irreversibly damaged, no longer needed, or are a threat to the organism can be eliminated quickly and neatly. In most cases, these deaths occur by apoptosis: the cells shrink, condense, and frequently fragment, and neighboring cells or macrophages rapidly phagocytose the cells or fragments before there is any leakage of cytoplasmic contents. Apoptosis depends on proteolytic enzymes called caspases, which cleave specific intracellular proteins to help kill the cell. Caspases are present in all nucleated animal cells as inactive precursors called procaspases. Initiator procaspases are activated when brought into proximity in activation complexes: once activated, they cleave and activate downstream executioner procaspases, which activate other executioner procaspases (and various other target proteins in the cell), producing an amplifying, irreversible proteolytic cascade.

Cells use at least two distinct pathways to activate initiator procaspases and trigger a caspase cascade leading to apoptosis: the extrinsic pathway is activated by extracellular ligands binding to cell-surface death receptors; the intrinsic pathway is activated by intracellular signals generated when cells are stressed. Each pathway uses its own initiator procaspases, which are activated in distinct activation complexes, called the DISC and the apoptosome, respectively. In the extrinsic pathway, the death receptors recruit procaspases-8 and 10 via adaptor proteins to form the DISC; in the intrinsic pathway, cytochrome c released from the intermembrane space of mitochondria activates Apaf1, which assembles into an apoptosome and recruits and activates procaspase-9.

Both extracellular signal proteins and intracellular Bcl2 proteins and IAP proteins tightly regulate the apoptotic program to ensure that cells normally kill themselves only when it benefits the animal. Both anti-apoptotic and pro-apoptotic Bcl2 proteins regulate the intrinsic pathway by controlling the release of mitochondrial intermembrane proteins, while IAP proteins inhibit activated caspases and promote their degradation.

PROBLEMS

Which statements are true? Explain why or why not.

18–1 In normal adult tissues, cell death usually balances cell division.

18–2 Mammalian cells that do not have cytochrome *c* should be resistant to apoptosis induced by UV light.

Discuss the following problems.

18–3 One important role of Fas and Fas ligand is to mediate elimination of tumor cells by killer lymphocytes. In a study of 35 primary lung and colon tumors, half the tumors were found to have amplified and overexpressed a gene for a secreted protein that binds to Fas ligand. How do you suppose that overexpression of this protein might contribute to the survival of these tumor cells? Explain your reasoning.

18–4 Development of the nematode *Caenorhabditis elegans* generates exactly 959 somatic cells; it also produces an additional 131 cells that are later eliminated by programmed cell death. Classical genetic experiments in *C. elegans* isolated mutants that identified the first genes involved in apoptosis. Of the many mutant genes affecting apoptosis in the nematode, none have ever been found in the gene for cytochrome *c*. Why do you suppose that such a central effector molecule in apoptosis was not found in the many genetic screens for "death" genes that have been carried out in *C. elegans*?

18–5 Imagine that you could microinject cytochrome *c* into the cytosol of wild-type cells and of cells that were doubly defective for Bax and Bak. Would you expect one, both, or neither type of cell to undergo apoptosis? Explain your reasoning.

18–6 In contrast to their similar brain abnormalities, newborn mice deficient in Apaf1 or caspase-9 have distinctive abnormalities in their paws. Apaf1-deficient mice fail to eliminate the webs between their developing digits, whereas caspase-9-deficient mice have normally formed digits (**Figure Q18–1**). If Apaf1 and caspase-9 function in the same apoptotic pathway, how is it possible for these deficient mice to differ in web-cell apoptosis?

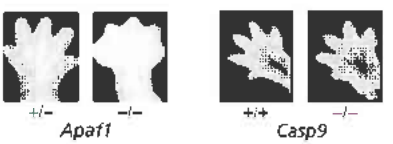


Figure Q18–1 Appearance of paws in *Apaf1*$^{-/-}$ and *Casp9*$^{-/-}$ newborn mice relative to normal newborn mice (Problem 18–6). (From H. Yoshida et al., *Cell* 94:739–750, 1998. With permission from Elsevier.)

18–7 When human cancer (HeLa) cells are exposed to UV light at 90 mJ/cm^2, most of the cells undergo apoptosis within 24 hours. Release of cytochrome *c* from mitochondria can be detected as early as 6 hours after exposure of a population of cells to UV light, and it continues to increase for more than 10 hours thereafter. Does this mean that individual cells slowly release their cytochrome *c* over this time period? Or, alternatively, do individual cells release their cytochrome *c* rapidly but with different cells being triggered over the longer time period?

To answer this fundamental question, you have fused the gene for green fluorescent protein (GFP) to the gene for cytochrome *c*, so that you can observe the behavior of individual cells by confocal fluorescence microscopy. In cells that are expressing the cytochrome *c*–GFP fusion, fluorescence shows the punctate pattern typical of mitochondrial proteins. You then irradiate these cells with UV light and observe individual cells for changes in the punctate pattern. Two such cells (outlined in white) are shown in **Figure Q18–2A** and B. Release of cytochrome *c*–GFP is detected as a change from a punctate to a diffuse pattern of fluorescence. Times after UV exposure are indicated as hours:minutes below the individual panels.

Which model for cytochrome c release do these observations support? Explain your reasoning.

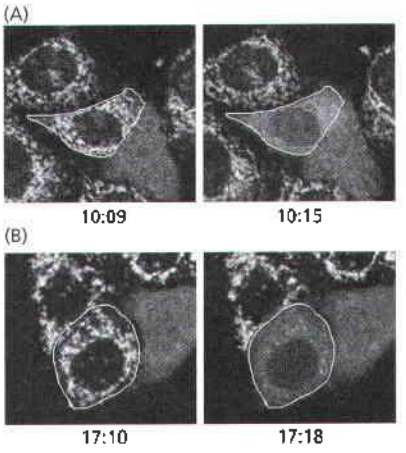


Figure Q18–2 Time-lapse video, fluorescence microscopic analysis of cytochrome *c*–GFP release from mitochondria of individual cells (Problem 18–7). (A) Cells observed for 8 minutes, 10 hours after UV irradiation. (B) Cells observed for 6 minutes, 17 hours after UV irradiation. One cell in (A) and one in (B), each *outlined in white*, have released their cytochrome *c*–GFP during the time frame of the observation, which is shown as hours:minutes below each panel. (From J.C. Goldstein et al., *Nat. Cell Biol.* 2:156–162, 2000. With permission from Macmillan Publishers Ltd.)

REFERENCES

Adams JM, Huang DC, Strasser A et al (2005) Subversion of the Bcl-2 life/death switch in cancer development and therapy. *Cold Spring Harb Symp Quant Biol* 70:469–77.

Boatright KM & Salvesen GS (2003) Mechanisms of caspase activation. *Curr Opin Cell Biol* 15:725–731.

Danial NN & Korsmeyer SJ (2004) Cell death: critical control points. *Cell* 116:205–219.

Ellis RE, Yuan JY & Horvitz RA (1991) Mechanisms and functions of cell death. *Annu Rev Cell Biol* 7:663–698.

Fadok VA & Henson PM (2003) Apoptosis: giving phosphatidylserine recognition an assist—with a twist. *Curr Biol* 13:R655–R657.

Galonek HL & Hardwick JM (2006) Upgrading the BCL-2 network. *Nature Cell Biol* 8:1317–1319.

Green DR (2005) Apoptotic pathways: ten minutes to dead. *Cell* 121:671–674.

Horvitz HR (2003) Worms, life, and death (Nobel lecture). *Chembiochem* 4:697–711.

Hyun-Eui K, Fenghe D, Fang M & Wang X (2005) Formation of apoptosome is initiated by cytochrome c-induced dATP hydrolysis and subsequent nucleotide exchange on Apaf-1. *Proc Natl Acad Sci USA* 102:17545–17550.

Jacobson MD, Weil M & Raff MC (1997) Programmed cell death in animal development. *Cell* 88:347–354.

Jiang X & Wang X (2004) Cytochrome C-mediated apoptosis. *Annu Rev Biochem* 73:87–106.

Kerr JF, Wyllie AH & Currie AR (1972) Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer* 26:239–257.

Kumar S (2007) Caspase function in programmed cell death. *Cell Death Differ* 14:32–43.

Lavrik I, Golks A & Krammer PH (2005) Death receptor signaling. *J Cell Sci* 118:265–267.

Lowe SW, Cepero E & Evan G (2004) Intrinsic tumour suppression. *Nature* 432:307–315.

Lum JJ, Bauer DE, Kong M et al (2005) Growth factor regulation of autophagy and cell survival in the absence of apoptosis. *Cell* 120:237–48.

McCall K & Steller H (1997) Facing death in the fly: genetic analysis of apoptosis in *Drosophila*. *Trends Genet* 13:222–226.

Nagata S (1999) Fas ligand-induced apoptosis. *Annu Rev Genet* 33:29–55.

Nagata S (2005) DNA degradation in development and programmed cell death. *Annu Rev Immunol* 23:853–875.

Pop C, Timmer J, Sperandio S & Salvesen GS (2006) The apoptosome activates caspase-9 by dimerization. *Mol Cell* 22:269–275.

Raff MC (1999) Cell suicide for beginners. *Nature* 396:119–122.

Rathmell JC & Thompson CB (2002) Pathways of apoptosis in lymphocyte development, homeostasis, and disease. *Cell* 109:S97–107.

Tittel JN & Steller H (2000) A comparison of programmed cell death between species. *Genome Biol* 1.

Verhagen AM, Coulson EJ & Vaux DL (2001) Inhibitor of apoptosis proteins and their relatives: IAPs and other BIRPs. *Genome Biol* 2:3009.1–3009.10.

Vousden KH (2005) Apoptosis. p53 and PUMA: a deadly duo. *Science* 309:1685–1686.

Willis SN & Adams JM (2005) Life in the balance: how BH3-only proteins induce apoptosis. *Curr Opin Cell Biol* 17:617–625.