CHAPTER TEN

Structure-Based Drug Design

LARRY W. HARDY
Aurigene Discovery Technologies
Lexington, Massachusetts

DONALD J. ABRAHAM
Virginia Commonwealth University
Richmond, Virginia

MARTIN K. SAFO
Virginia Commonwealth University
Richmond, Virginia

Contents

1 Introduction, 418
2 Structure-Based Drug Design, 419
2.1 Theory and Methods, 419
2.2 Hemoglobin, One of the First Drug-Design Targets, 419
2.2.1 History, 419
2.2.2 Sickle-Cell Anemia, 419
2.2.3 Allosteric Effectors, 421
2.2.4 Crosslinking Agents, 424
2.3 Antifolate Targets, 425
2.3.1 Dihydrofolate Reductase, 425
2.3.2 Thymidylate Synthase, 426
2.3.2.1 Structure-Guided Optimization: AG85 and AG337, 426
2.3.2.2 De Novo Lead Generation: AG331, 428
2.3.3 Glycinamide Ribonucleotide Formyltransferase, 429
2.4 Proteases, 432
2.4.1 Angiotensin-Converting Enzyme and the Discovery of Captopril, 432
2.4.2 HIV Protease, 433
2.4.3 Thrombin, 442
2.4.4 Caspase-1, 444
2.4.5 Matrix Metalloproteases, 445
2.5 Oxidoreductases, 446
2.5.1 Inosine Monophosphate Dehydrogenase, 447
2.5.2 Aldose Reductase, 448
2.6 Hydrolases, 449
2.6.1 Acetylcholinesterase, 449
2.6.2 Neuraminidase, 450

Burger's Medicinal Chemistry and Drug Discovery
Sixth Edition, Volume 1: Drug Discovery
Edited by Donald J. Abraham
ISBN 0-471-27090-3 © 2003 John Wiley & Sons, Inc.

2.6.3 Phospholipase A2 (Nonpancreatic, Secretory), 452
2.7 Picornavirus Uncoating, 454
2.8 Phosphoryl Transferases, 456
2.8.1 Mitogen-Activated Protein Kinase p38α, 456
2.8.2 Purine Nucleoside Phosphorylase, 459
2.9 Conclusions and Lessons Learned, 461

1 INTRODUCTION

Structure-based drug design by use of structural biology remains one of the most logical and aesthetically pleasing approaches in drug discovery paradigms. The first paper on the potential use of crystallography in medicinal chemistry was written in 1974 (1) and was presented at Professor Alfred Burger's retirement symposium in 1972. The excerpted last paragraph in the paper, reproduced below, foresaw the integration of X-ray crystallography into the field of medicinal chemistry.

> It is reasonable to assume then that the future of large molecule crystallography in medical chemistry may well be of monumental proportions. The reactivity of the receptor certainly lies in the nature of the environment and position of various amino acid residues. When the structured knowledge of the binding capabilities of the active site residues to specify groups on the agonist or antagonists becomes known, it should lead to proposals for syntheses of very specific agents with a high probability of biological action. Combined with what is known about transyort of drugs through a Hansch-type analysis, etc., it is feasible that the drugs of the future will be tailormade in this fashion. Certainly, and unfortunately, however, this day is not as close as one would like. The X-ray technique for large molecules, crystallization techniques, isolation techniques of biological systems, mechanism studies of active sites and synthetic talents have not been extensively intertwined because of the existing barriers (1).

Since that time there have been numerous successes in advancing new agents into clinical trials by combining crystallography with associated fields in drug discovery. Currently, more structures are solved every year than were in the entire Protein Data Bank in 1972. Although almost every major pharmaceutical company has an X-ray diffraction group, Agouron (now Pfizer) was the first biotechnology startup company to make drug discovery based on X-ray crystallography a central and primary theme (2). Other startup companies (such as BioCryst and Vertex) were soon founded to apply similar approaches. More recent companies, such as Structural Genomix (3) and Astex (4), and the High Throughput Crystallography Consortium, organized by Accelrys (5), have emerged to carry on structure-based drug discovery in a high throughput environment. Third-generation synchrotron sources, such as the Advanced Photon Source (APS) at Argonne National Laboratory outside Chicago, and new detectors, have enormously increased the speed of data collection. It is now possible to collect high resolution data from protein crystals, solve, and refine the structure in days to a few months. This information is covered in an adjacent chapter. Simultaneous advances in computing have added to the speed of obtaining three-dimensional structural information on interesting drug design targets. These developments, coupled with the sequencing of the human genome and the advent of bioinformatics, provide workers in structure-based drug design with a plethora of opportunities for success.

The utility of any drug discovery tool is measured, in the final analysis, by the output of the tool's use. New tools are burdened with unrealistically high expectations. As their application begins, the impact is sometimes more limited than was originally envisaged. Structure-based design methods have had great utility for the design of enzyme inhibitors, tight-binding receptor ligands, and novel proteins. The utility of these methods for the design of ***drugs*** is somewhat more limited, simply because there are so many factors that must be balanced in the successful design of a drug. Nonetheless, structure-based drug design (SBDD), distinct from the (far easier) structure-based ***inhibitor*** design, is now a reality and has had significant impact. Aspects of the methods and utility of SBDD have been described in several excellent review articles and monographs (6–12). This chapter focuses on the utility of SBDD in the cases of drugs

that have been launched as products, or that have at least entered human clinical trials. In some cases, SBDD has been a remarkable suecess. In others, it has failed in the sense that, despite its use, the candidate produced did not gain approval to become a marketed drug. In the latter cases, this was usually not truly a failure of SBDD, but rather attributed to the complex criteria that drugs must meet, and to the complex regulatory hurdles that candidates and companies face.

In addition to providing a measure of the impact of SBDD on the creation of actual drugs, these examples will also provide lessons about how to apply SBDD in drug discovery. The chapter is not completely encyclopedic, and some significant instances of SBDD will be missed by the informed reader. However, the discovery programs with drugs and drug candidates that are discussed will provide sufficient diversity that general trends can emerge. In a few cases, relevant results for preclinical compounds that seem likely to enter human trials are described. A growing number of the drugs to which structural design methods are applied are themselves proteins (e.g., cytokines, immunomodulators, monoclonal antibodies). However, this chapter is restricted to small organic molecules that are designed by use of the three-dimensional structure of a target protein.

2 STRUCTURE-BASED DRUG DESIGN

2.1 Theory and Methods

The concept of structure-based drug discovery combines information from several fields: X-ray crystallography and/or NMR, molecular modeling, synthetic organic chemistry, qualitative structure–activity relationships (QSAR), and biological evaluation. Figure 10.1 represents a general road map where a cyclic process refines each stage of discovery. Initial binding site information is gained from the three-dimensional solution of a complex of the target with a lead compound(s). Molecular modeling is usually next applied with the intent of designing a more specific ligand(s) with higher affinity. Synthesis and subsequent in *vitro* biological evaluation of the new agents produces more candidates for crystallographic or NMR analysis, with the hope of correlating the biological action with precise structural information. It makes good sense at the early stages of design to use lead molecule structural scaffolds that retain low toxicity profiles, given that the latter most often derails successful drug discovery. The most active derivative(~) from this cyclic process can be forwarded for in vivo evaluation in animals.

2.2 Hemoglobin, One of the First Drug-Design Targets

2.2.1 History. Perutz and colleagues determined the first three-dimensional structures of proteins. Through use of X-ray crystallography Kendrew determined the structure of myoglobin (13), whereas Perutz determined the structure of hemoglobin (Hb) (14–16). At the present time, new protein and nucleic acid structures and complexes are published weekly. However, for a long period after the first protein structures were solved, progress was slower. Hb was of interest for drug discovery purposes because of the early identification of the mutant 6 Glu → Val, which causes sickle-cell anemia. The crystal structure of sickle Hb (Hbs) was published by Wishner et al. (17) and was later solved at a higher resolution by Harrington et al. (18).

2.2.2 Sickle-Cell Anemia. In 1975, through use of the three-dimensional Hb coordinates, two groups initiated SBDD studies to discover an agent to treat sickle-cell anemia: Goodford et al. in England and Abraham et al. in the United States. Goodford's group was the first to develop an antisickling agent (BW12C), based on structure-based drug design, which reached clinical trials (19, 20). However, Wireko and coworkers were unable to confirm the BW12C binding site proposed by Goodford (21). The second antisickling agent proposed by Abraham et al. to advance to clinical trials was the food additive vanillin (compound 1a) (22). The crystallographic binding site of BW12C (1b) was found to be at the N-terminal amino groups of the a-chains (21), whereas that of vanillin shows binding close to αHis103 and also at a minor site between βHis117 and βHis117 (22). A recently redetermined binding site of vanillin at a higher resolution shows weak binding to the N-terminal amino group

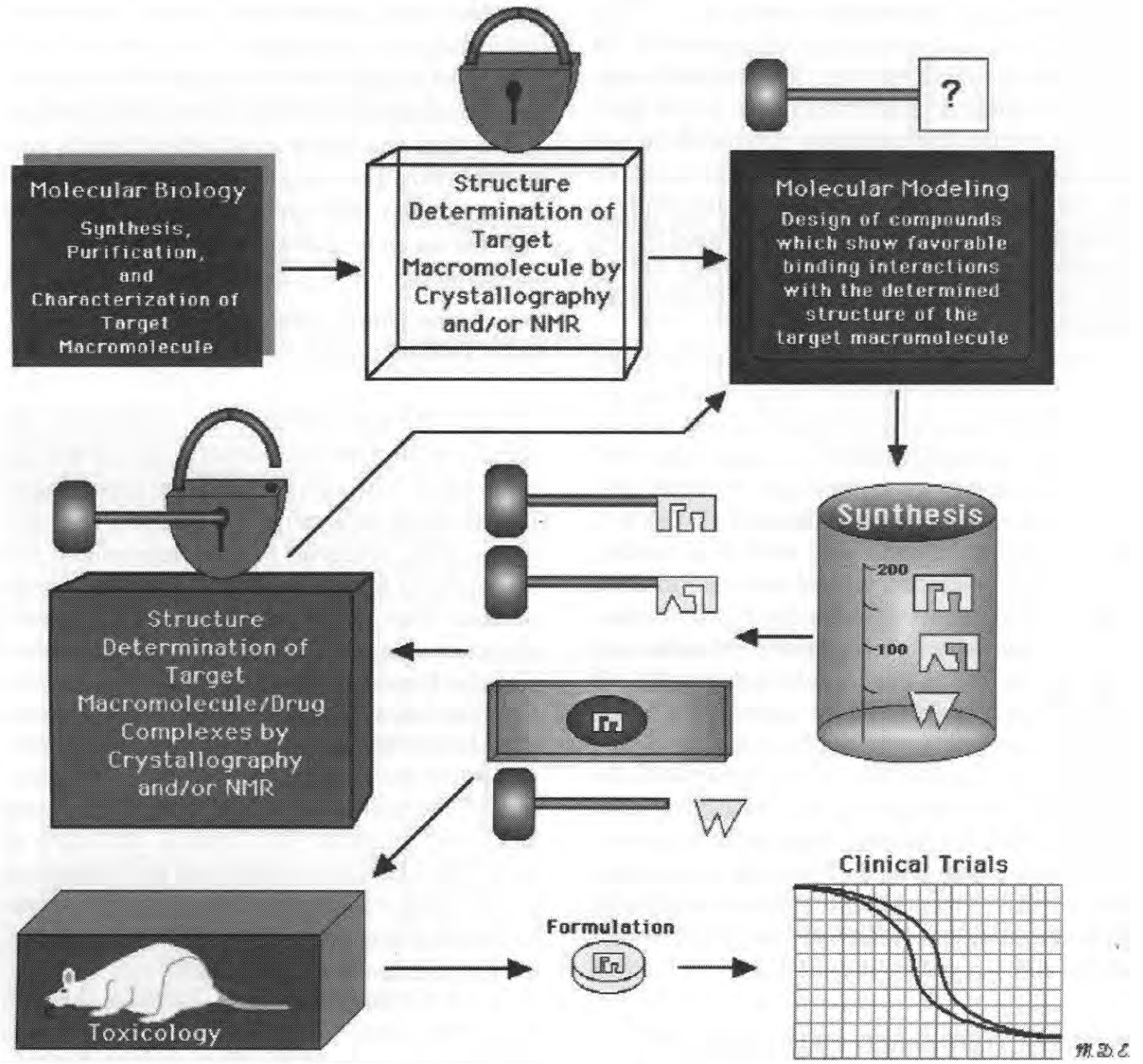
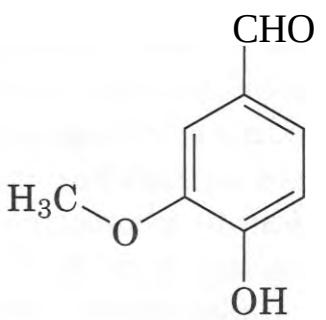
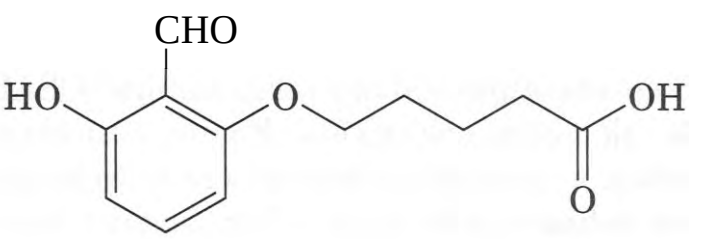


Figure 10.1. Schematic of the structure-based drug discovery/design process. The figure maps out the iterative steps that make use of X-ray crystallography, molecular modeling, organic synthesis, and biological testing to identify and optimize ligand–protein interactions.



(1a) vanillin



(Ib) BW12C

of the a-chain (23). A derivative of vanillin has been patented and is a candidate for clinical trials.

Two marketed medicines, ethacrynic acid (2), a diuretic agent, and clofibric acid (3), an antilipidemic agent, were reported to have strong antigelling activity (24, 25), and through X-ray analyses of cocrystals, the binding sites of these agents to Hb were elucidated (26). Unfortunately, it was found that high

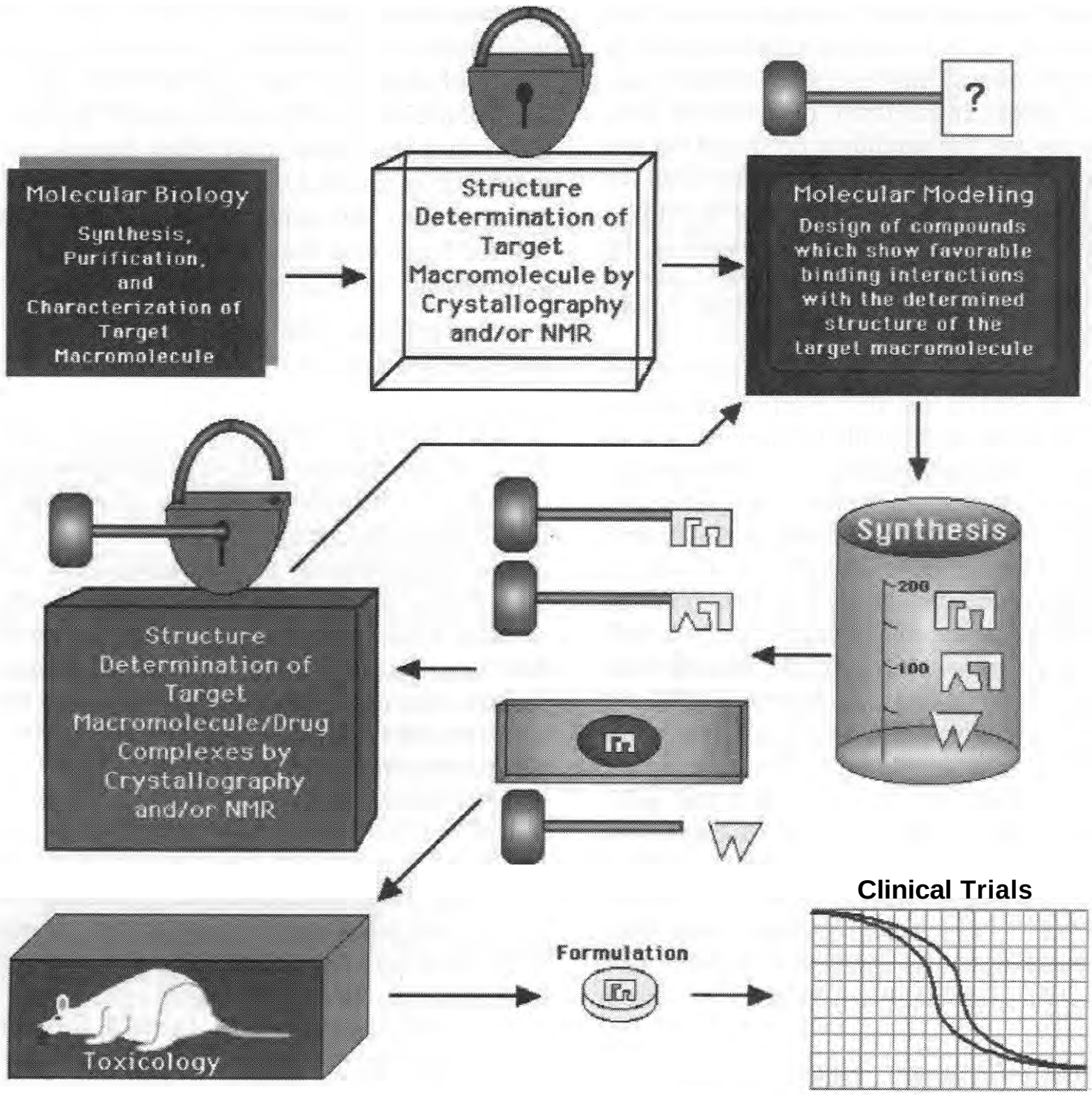
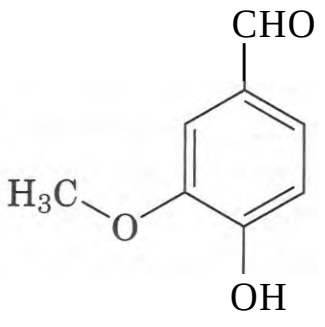
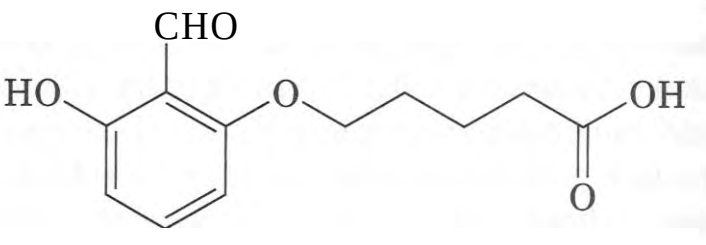


Figure 10.1. Schematic of the structure-based drug discovery/design process. The figure maps out the iterative steps that make use of X-ray crystallography, molecular modeling, organic synthesis, and biological testing to identify and optimize ligand–protein interactions.



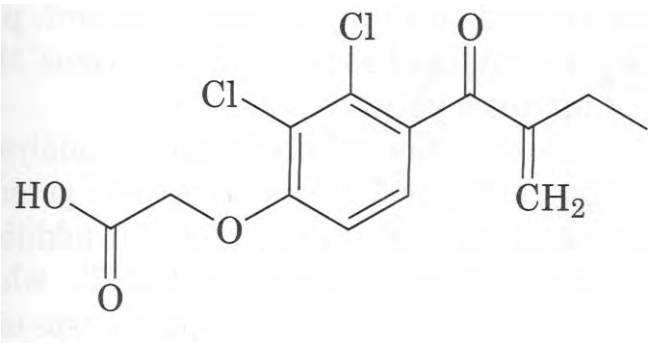
(1a) vanillin



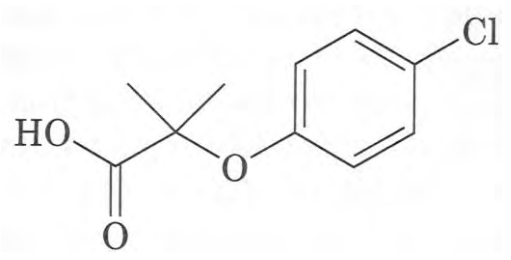
(1b) BW12C

of the a-chain (23). A derivative of vanillin has been patented and is a candidate for clinical trials.

Two marketed medicines, ethacrynic acid (2), a diuretic agent, and clofibric acid (3), an antilipidemic agent, were reported to have strong antigelling activity (24, 25), and through X-ray analyses of cocrystals, the binding sites of these agents to Hb were elucidated (26). Unfortunately, it was found that high



(2) ethacrynic acid



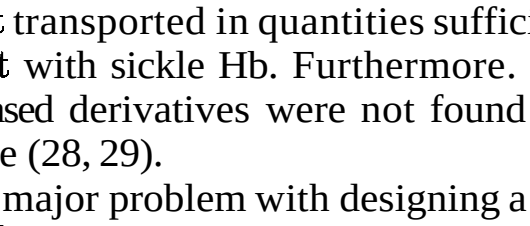
(3) clofibric acid

concentrations of ethacrynic acid were needed to interact with Hb in deformed red cells (27). Clofibric acid, when administered in a 2 gm/day dose (as the ethyl ester clofibrate), appeared to be an ideal potential treatment for sickle-cell anemia, but was subsequently found to be highly bound to serum proteins and not transported in quantities sufficient to interact with sickle Hb. Furthermore. structure-based derivatives were not found to be effective (28, 29).

The major problem with designing a small molecule to treat sickle cell anemia is not so much an issue of specificity, but arises from the treatment of a chronic disease. The potential cumulative toxicity from the amount of drug needed to interact with approximately two pounds of hemoglobin S over a homozygous patient's lifetime is the major concern (22) (for a review, see Vol. 3, Chapter 10. *Sickle* ***Cell Anemia,*** by Alan Schecter et al).

2.2!.3 *Allosteric* Effectors. *2,3-Diphospho*-glycerate (2,3-DPG, compound 4), found in most mammalian red cells, is the naturally occurring allosteric effector for Hb. Its physiological role is to right shift the Hb oxygen-binding curve to release more oxygen. The binding site of 2,3-DPG, determined by Arnone (30) lies on the dyad axis at the mouth of the β-cleft (Fig. 10.2) interacting with the N-terminal βVal1, βLys82, and βHis143 of deoxy Hb. A more recent study at a higher resolution, by Richard et al. (31), found DPG to interact with the residues βHis2 and βLys82. Goodford and colleagues were the first to design agents that would bind to the 2,3-DPG site (32–34). An effective allosteric effector that can enter red cells might be used to treat hypoxic diseases such as angina and stroke, to enhance radiation treatment of hypoxic tumors, or to extend the shelf life of stored blood.

Many antigelling agents left shift the oxygen binding curve, producing higher concentrations of oxy-HbS. Given to patients with sickle-cell anemia, this should result in less polymerization, and therefore less red blood cell sickling. It was a surprise therefore when clofibric acid, which blocks sickle-cell Hb polymerization, was found to shift the Hb oxygen binding curve to the right, in a manner similar to that of 2,3-DPG (25). The clofibric acid binding site was found to be far removed from the 2,3-DPG site (25, 35). The determination of the clofibric acid binding site on Hb was the first report of a tense state (deoxy state) allo-



(4) 2,3-DPG

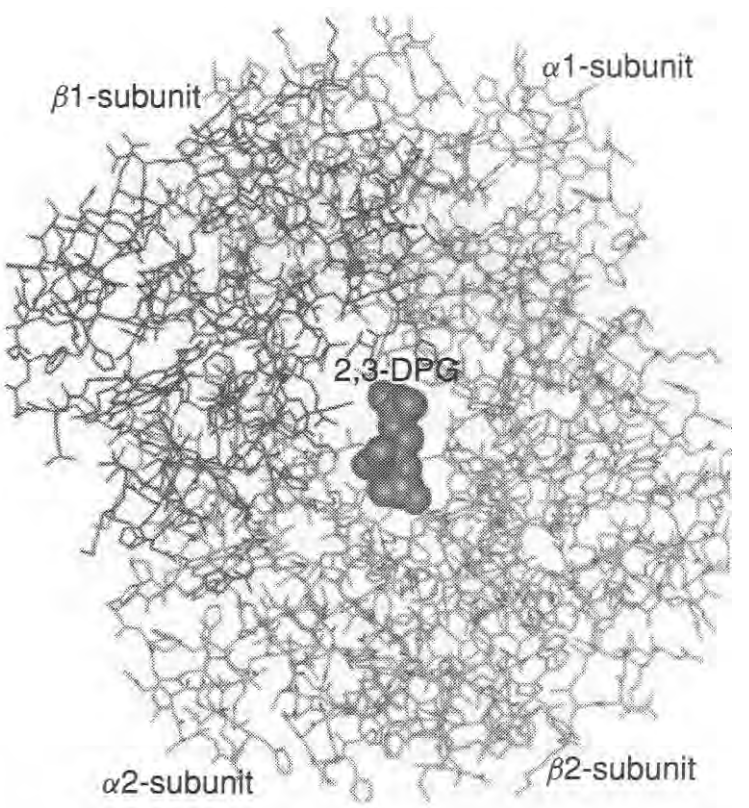
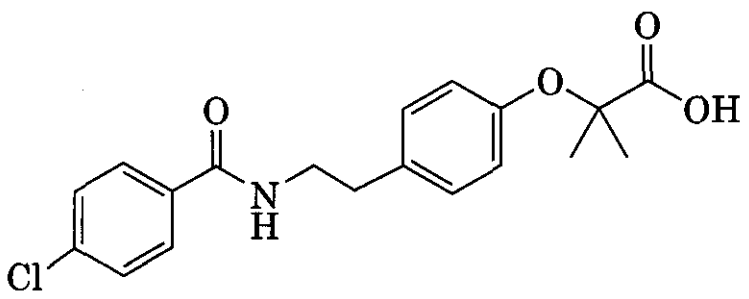


Figure 10.2. View of (4)(2,3-DPG) binding site at the mouth of the β-cleft of deoxy hemoglobin. See color insert.

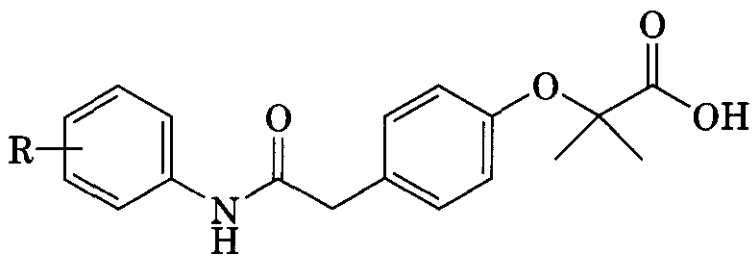
steric binding site different from that of 2,3-DPG (compound 4). Perutz and Poyart tested another antilipidemic agent, bezafibrate (compound 5), and found that it was an even more



(5) bezafibrate

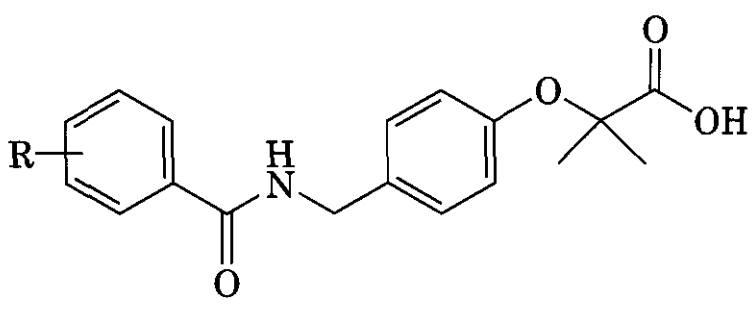
potent right-shifting agent than clofibrate (36). Perutz et al. (26) and Abraham (35) determined the binding site of bezafibrate and found it to link a high occupancy clofibrate site with a low occupancy site. Lalezari and Lalezari synthesized urea derivatives of bezafibrate (37), and with Perutz et al. determined the binding site of the most potent derivatives (38). Although these compounds were extremely potent, they were hampered by serum albumin binding (39, 40).

Abraham and coworkers synthesized a series of bezafibrate analogs (39–42). One of these agents, efaproxaril (RSR 13, compound 6a) is currently in Phase III clinical trials for radiation treatment of metastatic brain tumors (see, Vol. 4, Chapter 4. ***Radiosensitizers and Radioprotective Agents,*** by Edward Bump et al). The binding constants and binding sites of a large number of these bezafibrate analogs were measured and agreed with the number of crystallographic binding sites found (42). The degree of right shift in the oxygen-binding curve produced by these compounds was not solely related to their binding constant, providing a structural basis for E. J. Ariens' theory of intrinsic activity (42).



(6a) (RSR-13) R = 3,5-dimethyl
(6b) (RSR-56) R = 3,5-dimethoxy

By use of X-ray crystallographic analyses, the key elements linking allosteric potency with structure were uncovered. In addition, the computational program HINT, which quantitates atom–atom interactions, was used to determine the strongest contacts between various bezafibrate analogs and Hb residues. These analyses revealed that the amide linkage between the two aromatic rings of the compounds must be orientated so that the carbonyl oxygen forms a hydrogen bond with the side-chain amine of αLys99 (41, 43). Three other important interactions were found. The first are the water-mediated hydrogen bonds between the effector molecule and the protein, the most important occurring between the effector's terminal carboxylate and the side-chain guanidinium moiety of residue αArg141. Second, a hydrophobic interaction involves a methyl or halogen substituent on the effector's terminal aromatic ring and a hydrophobic groove created by Hb residues αPhe36, αLys99, αLeu100, αHis103, and βAsn108. Third, a hydrogen bond is formed between the side-chain amide nitrogen of Asn108 and the electron cloud of the effector's terminal aromatic ring (40, 41, 43). Abraham first observed this last interaction while elucidating the Hb binding site of bezafibrate (35). Burley and Petsko had previously pointed out this type of hydrogen bond in a number of proteins, indicating that this contact is involved in a number of other receptor interactions (44, 45). Perutz and Levitte estimated this bond to be about 3 kcal/mol (46). Figure 10.3 shows the overlap of four allosteric effectors (**6a**, 6b, 7a and 7b) that bind at the same site in deoxy Hb but differ in their allosteric potency.



(7a) (MM-30) R = 3,5-dichloro
(7b) (MM-25) R = 4-chloro

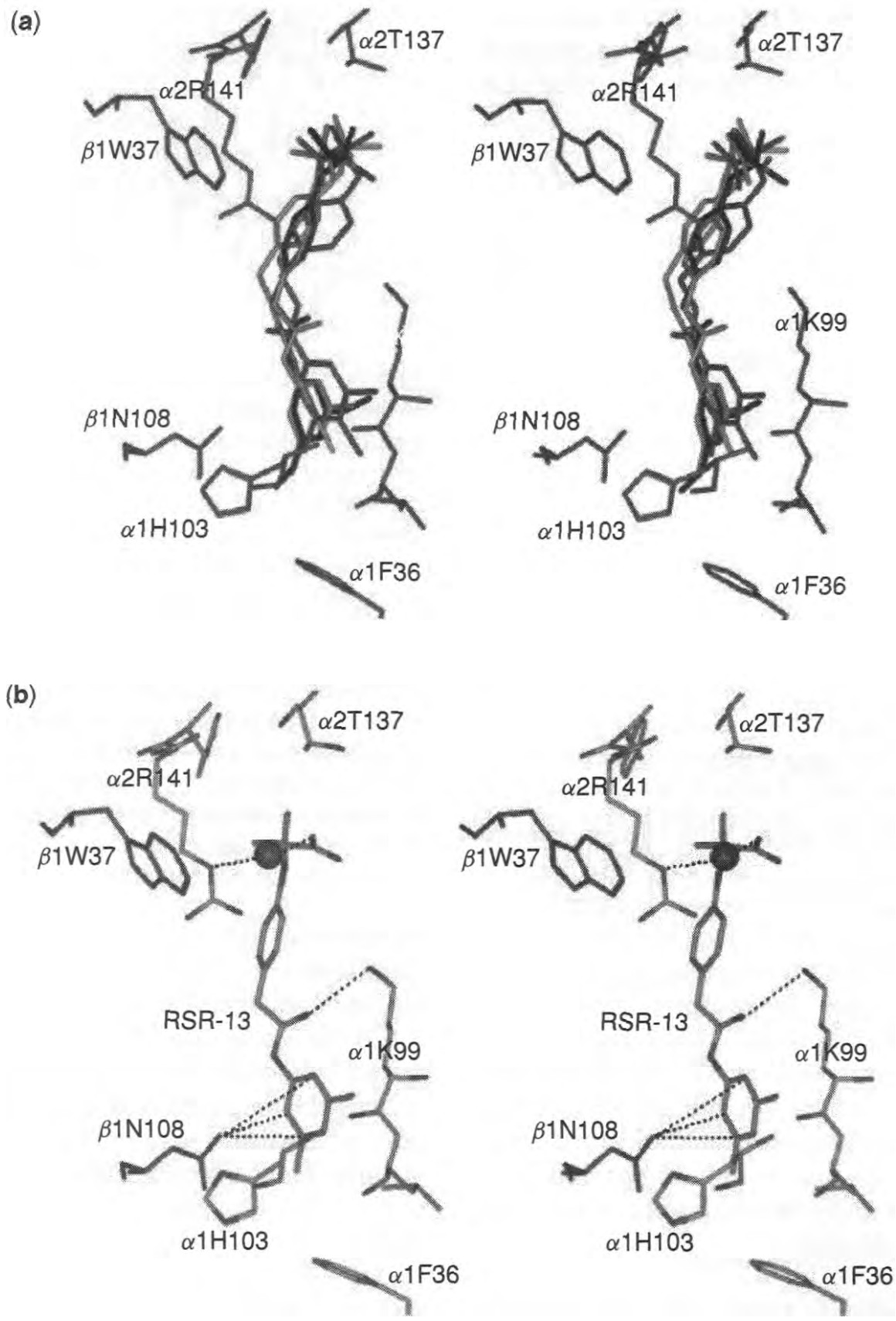
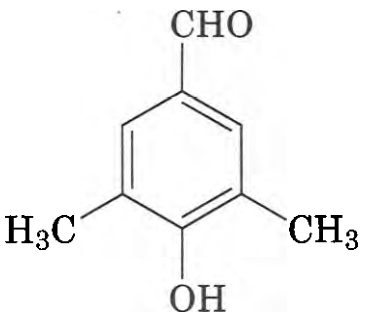
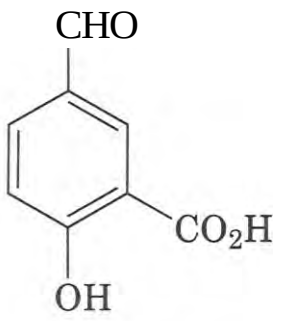


Figure **10.3.** Stereoview of allosteric binding site in deoxy hemoglobin. **A** similar compound environment is observed at the symmetry-related site, not shown here. (a) Overlap of four right-shifting allosteric effectors of hemoglobin: (**6a**) (RSR13, yellow), (**6b**) (RSR56, black), (**7a**) (MM30, red), and (**7b**) (MM25, cyan). The four effectors bind at the same site in deoxy hemoglobin. The stronger acting RSR compounds differ from the much weaker MM compounds by reversal of the amide bond located between the two phenyl rings. As a result, in both RSR13 and RSR56, the carbonyl oxygen faces and makes a key hydrogen bonding interaction with the amine of αLys99. In contrast, the carbonyl oxygen of the MM compounds is oriented away from the αLys99 amine. The αLys99 interaction with the RSR compounds appears to be critical in the allosteric differences. (b) Detailed interactions between RSR13 (**6a**) and hemoglobin, showing key hydrogen bonding interactions that help constrain the T-state and explain the allosteric nature of this compound and those of other related compounds. See color insert.

Over the course of these studies, an interesting anomaly was solved. Allosteric effectors (such as **8a** and **8b**) can bind to a similar site



(8a) DMHB



(8b) 5-FSA

and yet effect opposite shifts in the oxygen-binding curve. Agents such as 5-FSA bind to the N-terminal Val and provide groups for hydrogen bonding with the opposite dimer (across the twofold axis) right shift the oxygen-binding curve. In contrast, agents that disrupt the water-mediated linkage between the N-terminal αVal with the C-terminal αArg141 left shift the curve (47) (Fig. 10.4). Structure-based stereospecific allosteric effectors for Hb have also been developed and possess activities and profiles appropriate for clinical efficacy (48, 49).

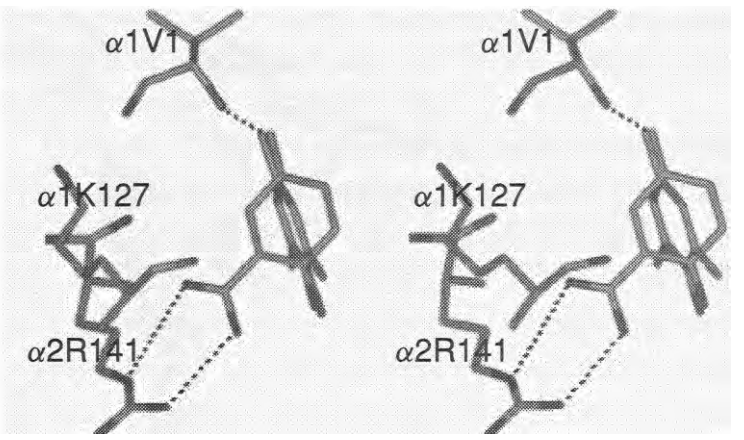
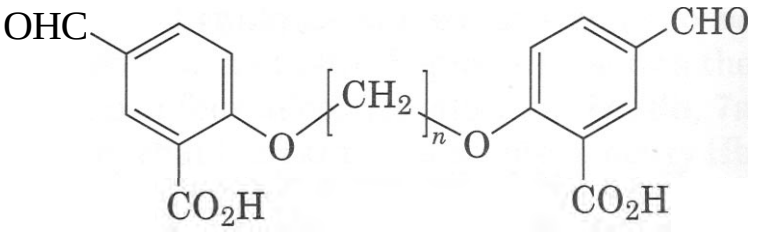


Figure 10.4. Stereoview of superimposed binding sites for (**8b**) (5-FSA, yellow) and (**8a**) (DMHB, magenta) in deoxy hemoglobin. A similar compound environment is observed at the symmetry-related site and therefore not shown here. Both compounds form a Schiff base adduct with the α1Val1 *N*-terminal nitrogen. Whereas the m-carboxylate of 5-FSA forms a salt bridge with the α2Arg141 (opposite subunit), this intersubunit bond is missing in DMHB. The added constraint to the T-state by 5-FSA that ties two subunits together shifts the allosteric equilibrium to the right. On the other hand, the binding of DMHB does not add to the T-state constraint. Instead, it disrupts any T-state salt- or water-bridge interactions between the opposite a-subunits. The result is a left shift of the oxygen equilibrium curve by DMHB. See color insert.

2.2.4 Crosslinking Agents. The first crosslinking agent that possessed potential as a Hb-based blood substitute was described by Walder et al. (50). Bis(4-formylbenzyl)ethane and bisulfite adducts of similar symmetrical aromatic dialdehydes, previously studied by Goodford and colleagues, provided the starting points that led to these compounds. Chatterjee et al. identified the binding site to deoxy-Hb, and found that the two Lys99 side chains were crosslinked (51). One of the derivatives was proposed as a blood substitute (52), and has been explored commercially (see Vol. 3, Chapter 8. Oxygen Delivery and Blood *Substitutes* and Blood Products, by Andeas Mozzarelli et al.). Another crosslinked Hb engineered by Nagai and colleagues, at the MRC-LMB in Cambridge, was developed into a blood substitute that was clinically investigated at Somatogen, now Baxter (53). Boyiri et al. synthesized a number of crosslinking agents (molecular ratchets, such as 9) whose



(9) n = 1–10 TB36, n = 3

potency was directly related to the length of the crosslink: the shorter the crosslink (three atoms), the stronger the shift of the oxygen binding curve to the right (54) (Fig. 10.5).

Perutz's hypothesis (55) and the MWC model (56) for allostery, that the more tension is added to the tense (deoxy) state of Hb, the greater the shift to the right of the oxygen-

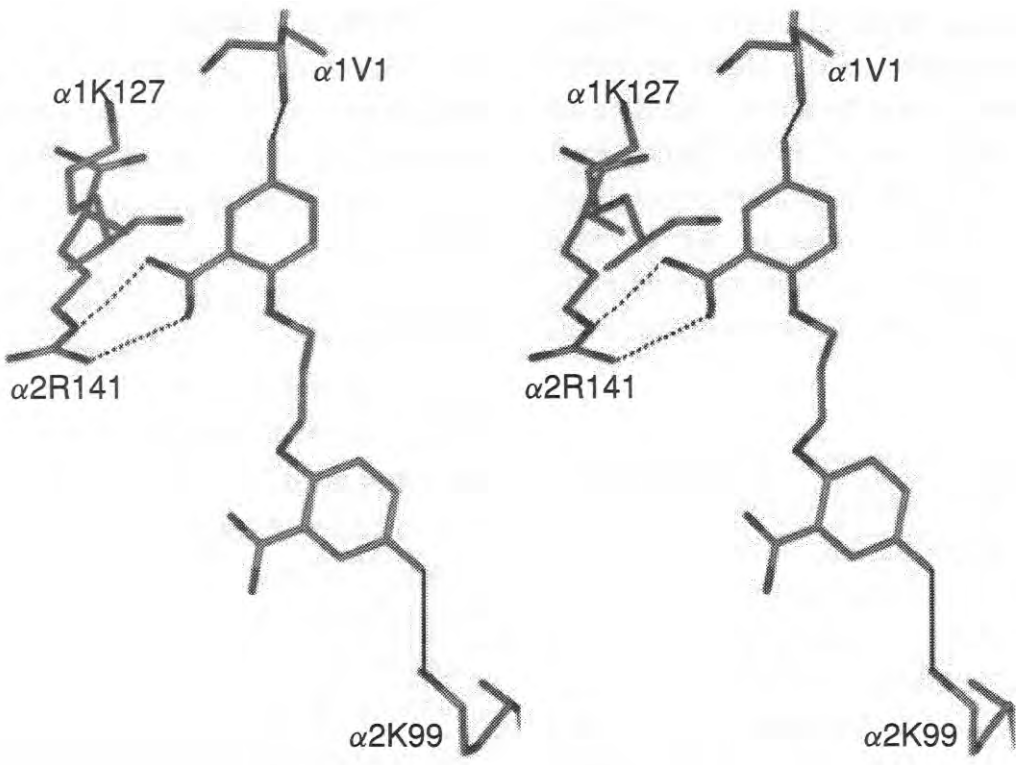
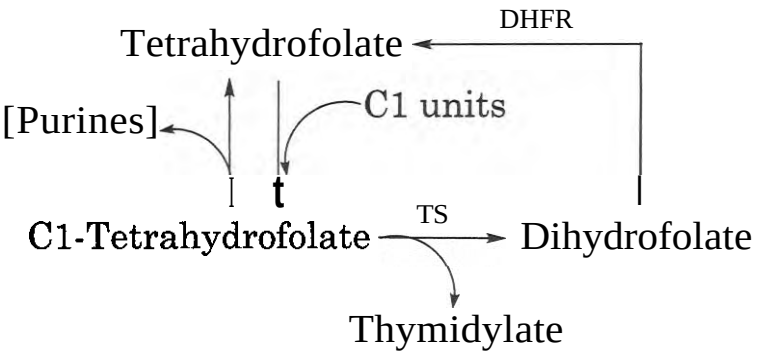


Figure 10.5. Stereoview of the binding site for (9) (n = 3, TB36, yellow) in deoxy Hb. A similar compound environment is observed at the symmetry-related site, not shown here. One aldehyde is covalently attached to the N-terminal α1Val1, whereas the second aldehyde is bound to the opposite subunit, α2Lys99 ammonium ion. The carboxylate on the first aromatic ring forms a bidentate hydrogen bond and salt bridge with the guanidinium ion of α2Arg141 of the opposite subunit. The effector thus ties two subunits together and adds additional constraints to the T-state, resulting in a shift in the Hb allosteric equilibrium to the right. The magnitude of constraint placed on the T-state by the crosslinked αLys99 varies with the flexibility of the linker. Shorter bridging chains form tighter crosslinks and yield larger shifts in the allosteric equilibrium. See color insert.

binding curve, are generally consistent with the behavior of the allosteric effectors and crosslinking agents.

2.3 Antifolate Targets

2.3.1 Dihydrofolate Reductase. The reduced form of folate (tetrahydrofolate) acts as a one-carbon donor in a wide variety of biosynthetic transformations. This includes essential steps in the synthesis of purine nucleotides and of thymidylate, essential precursors to DNA and RNA. For this reason. folate-dependent enzymes have been useful targets for the development of anticancer and anti-inflammatory drugs (e.g., methotrexate) and anti-infectives (trimethoprim, pyrimethamine). During the reaction catalyzed by thymidylate synthase (TS), tetrahydrofolate also acts as a reductant and is converted stoichiometrically to dihydrofolate. The regeneration of tetrahydrofolate, required for the continuous functioning of this cofactor, is catalyzed by dihydrofolate reductase (DHFR).



Scheme 10.1.

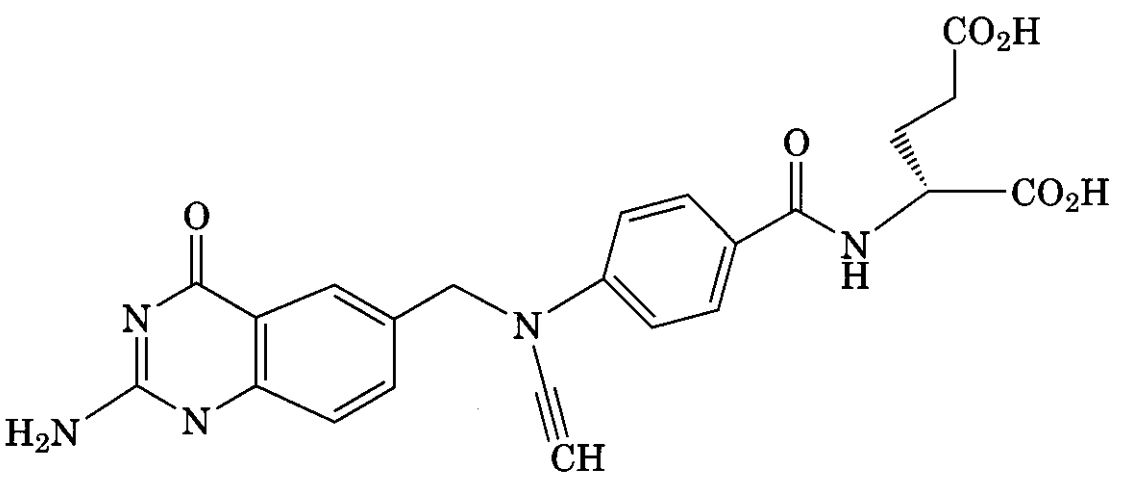
The first crystal structure of a drug bound to its molecular target was provided by the pioneering X-ray diffraction study of the complex between DHFR and methotrexate (57), albeit in this case the target was a bacterial surrogate for the actual target (the human enzyme). Once X-ray structures of DHFR from eukaryotic sources were also solved, comparisons of the bacterial and eukaryotic DHFR structures revealed the structural basis for the selectivity of the antibacterial drug trimethoprim for the bacterial enzyme. This understanding allowed Goodford and colleagues

to rationally design trimethoprim analogs with altered potencies (58). Retrospective studies such as those done by David Matthews and others on DHFR (see, for example, Ref. 59) set the stage for the iterative process of structure-based inhibitor design as it was later developed at Agouron Pharmaceuticals, targeted against another folate-dependent enzyme, TS (60, 61).

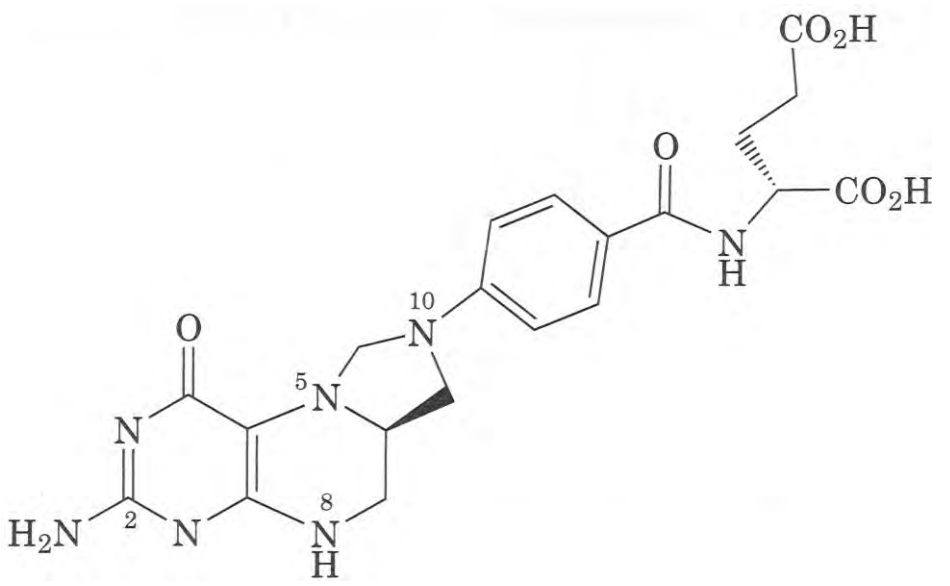
2.3.2 Thymidylate Synthase. There are two main modes in which structure-based methods for inhibitor design have been employed. The first mode is structure-guided optimization of the design of a previously known chemical scaffold. The scaffold could be a known drug or inhibitor, substrate analog, or a hit from screening of a random library. The property, which is modified during the optimization, may be, for example, potency, solubility, or target selectivity, or the more challenging aim may be to optimize several properties simultaneously. A second and potentially more powerful mode is for the ***de nouo*** design of inhibitory ligands, sometimes referred to as lead generation. This mode relies strictly on the structure of the target enzyme or receptor as a template. A substrate or an inhibitor may be bound to the crystalline target, and deleted to provide the template. This is advantageous when, as in the case of TS, a substantial conformational change occurs when ligands bind. After a ***de nouo*** design process has provided a new inhibitor that is structurally unique, the properties of the new scaffold can be optimized by continued structural guidance. Both modes of SBDD have been used to generate TS inhibitors that have entered clinical trials.

When the design of inhibitors of human TS at Agouron Pharmaceuticals began, the amounts of the human enzyme required for crystallographic study were unavailable. Because the active site of the enzyme is so highly conserved, it was assumed that an acceptable surrogate for human TS would be the crystal structure of a bacterial TS (60, 62). Figure 10.6 shows the conformation of the quinazoline folate analog 10 (N10-propynyl-5, 8-dideazafolate), bound within the active site of the ***Escherichia coli*** enzyme with the nucleotide substrate, 2′-deoxyuridine-5′-monophosphate (63, 64). This folate analog, designed by classical medicinal chemistry as an analog of the TS substrate, 5,10-methylene-tetrahydrofolate (**11**), is a potent TS inhibitor. Nevertheless, (10) failed as an anticancer drug because of its insolubility and resulting nephrotoxicity (65).

2.3.2.1 ***Structure-Guided Optimization: AG85 and AG337.*** In the crystalline complex with ***E. coli*** TS, the quinazoline ring of compound (10) binds on top of the pyrimidine of the nucleotide, in a protein crevice surrounded by hydrophobic residues (Fig. 10.6). The bound molecule bends at right angles between the quinazoline and 4-aminobenzoyl rings (at N10), with the D-glutamate portion extending out to the surface of the enzyme. Hydrogen bonds are made with several enzyme sidechains, the terminal carboxylate, and several tightly bound waters. This compound, like folate and most folate analogs, gains entry into cells through a transport system that recognizes its D-glutamate moiety, and intracellular concentrations are elevated because of trap-



(10) N10-propynyl-5,8-dideazafolate (also known as PDDF or CB3717)



(11) 5,10-methylene, 5,6,7,8-tetrahydrofolate

ping of the compound as highly charged forms after addition of several additional glutamates by a cellular enzyme.

TS inhibitors were designed by Agouron scientists with the aim of providing a drug that could enter cells passively and thus avoid the need for transport or polyglutamylation. The first were designed by structure-guided modification of known antifolates, and others were designed de *novo*. Starting with (**12**), the glutamlate moiety was deleted from the structure. [Compound (**12**), the 2-desamino-2-methyl analog of (**10**), had been found to be much more water soluble than (10). This eventually led (65) to AstraZeneca's Tomudex, which is now approved for treatment of colorectal cancer in European markets.] Removal of the glutamate reduced the potency by 2 to **3** orders of magnitude (Table 10.1, 12 versus 13). The crystal structure solved by use of (10) indicated potential interactions that were exploited by substituents such as the m-CF, in compound (14). The phenyl moiety in (15) was added to interact with Phe176 and Ile79 (Fig. 10.6). Combining substituents does not necessarily produce the expected sum of binding free energy (compare 16 with 14 and **15**). Structures of the complexes with several of these compounds revealed that ideal placement of one group does not always accommo-

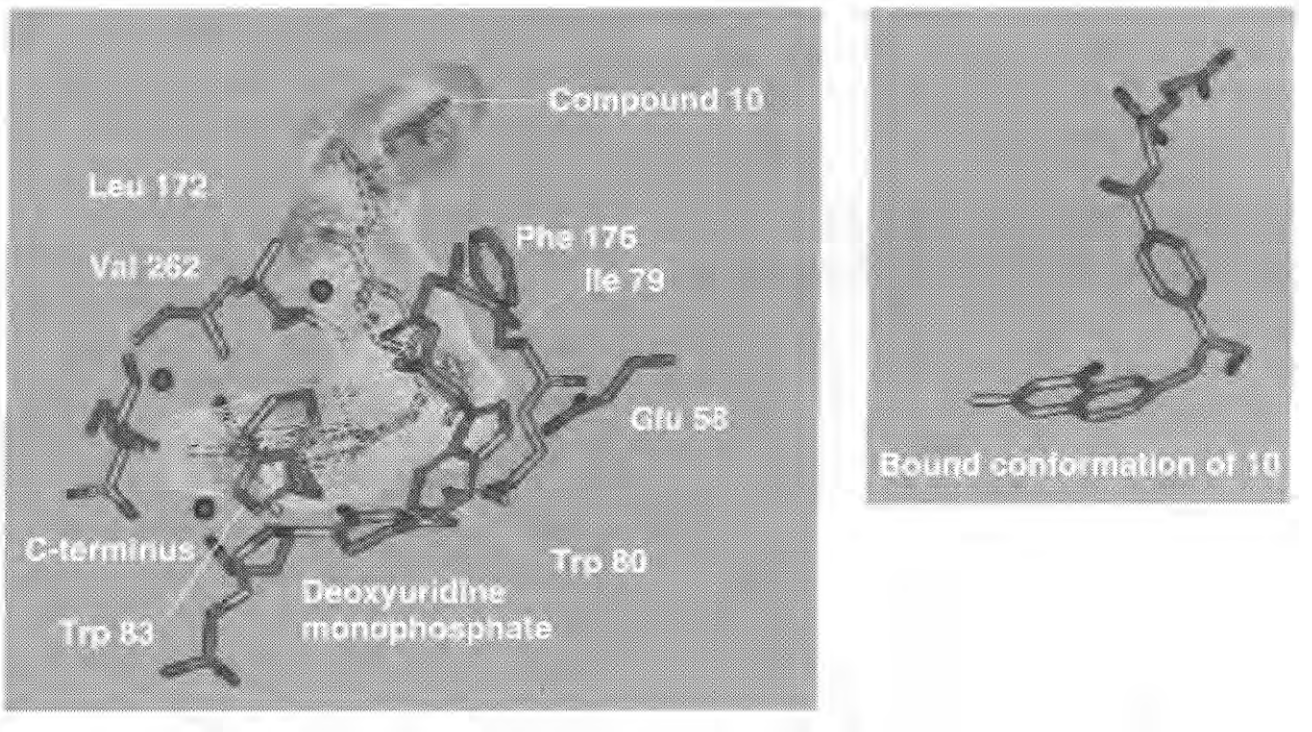
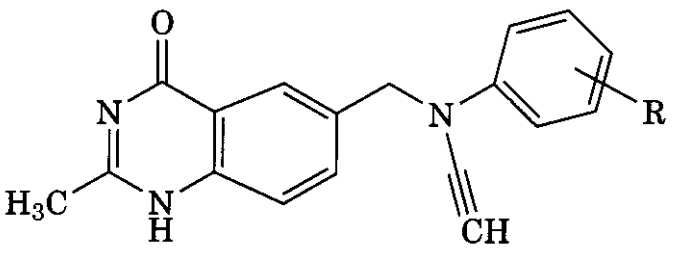


Figure 10.6. Binding site for **(10)** (*N10*-propynyl-5,8-dideazafolate), within the active site of thymidylate synthase from ***Escherichia coli.*** The surface of the inhibitor is shown in the left view. The red spheres in the left view are tightly bound water molecules. See color insert.

Table 10.1 **SAR for 2-Methyl-4-oxo-quinazoline Inhibitors of *TS*[a]**



(12)–(17)

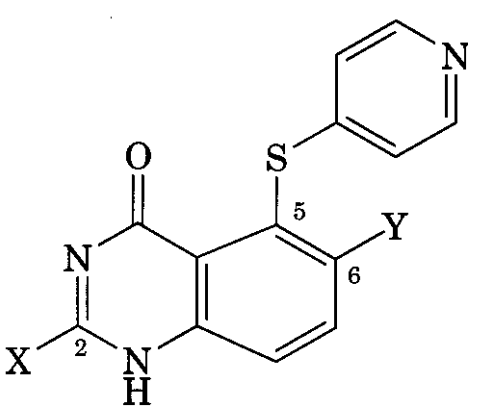
Compound	R	K_{is}, μM (*E. coli* TS)	K_{is}, μM (human TS)
(12)	para-CO-glutamate	0.005	0.009
(13)	H	4	2.2
(14)	*meta*-trifluoromethyl	0.45	0.4
(15)	para-SO,-phenyl	0.025	0.013
(16)	meta-trifluoromethyl, para-SO_2-phenyl	0.037	0.05
(17)	para-SO_2-(N-indolyl)	0.15	0.07

"From ref. 60.

date the best interaction for another. (This is a general problem for rigid scaffolds.) Compounds (**15–17**) had significant activity in ***in vitro*** cell-based assays, which could be reversed by exogenous thymidine. Compound (**17**) (AG85) was tested in human clinical trials for treatment of psoriasis (9).

The structure shown in Fig. 10.6 also suggested another approach to alter the structure of (12) to generate a lipophilic inhibitor of TS. The hydrophobic cavity filled by the aromatic ring of the para-aminobenzoyl group could be filled instead by a substituent attached to position 5 of the quinazoline nucleus. Four different 5-substituted 2-methyl-4-oxoquinazolines were made to test this idea, and one of these (**18**) was a 1 μM inhibitor of human TS (66).

The X-ray structure of the bacterial enzyme with (18) confirmed the hypothetical

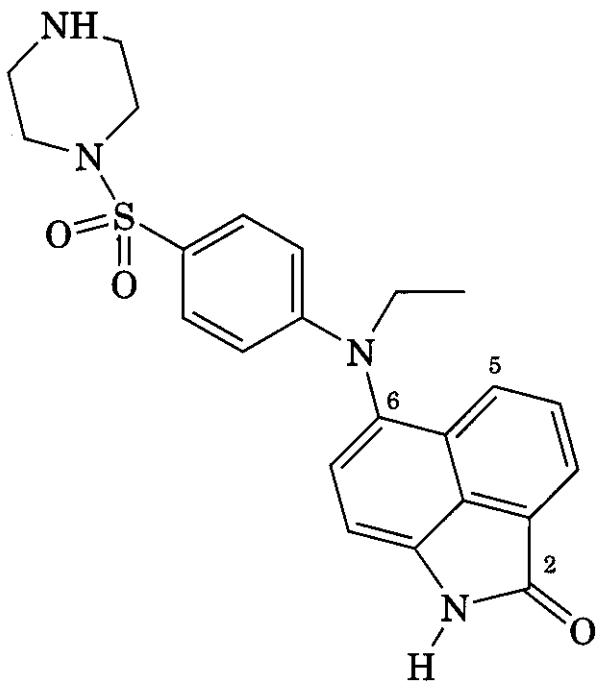


(18) X = CH_3, Y = H
(19) X = NH_2, Y = CH_3

binding mode. Two dozen 5-substituted quinazolines were made to explore the SAR for this scaffold. However, the eventual clinical candidate (**19**) was only two steps away from (**18**). The methyl group at position 6 was incorporated for favorable interaction with Trp80. This also favorably restricted the torsional flexibility for the 5-substituent, and increased the inhibitory potency against human TS by 10-fold. The 2-methyl was replaced by an amino group, to create a hydrogen bond to a backbone carbonyl in the protein, and increased potency another sixfold. Compound (19)(AG337, also known as nolatrexed, and as the hydrochloride, Thymitaq) advanced into human testing and had progressed into later-stage clinical trials **as** an antitumor agent by 1996 (67).

2.3.2.2 De Novo Lead Generation: AG331. The ***de novo*** design effort was initiated through the use of a computational method, Goodford's GRID algorithm (68, 69), to locate a site favorable to the binding of an aromatic system within the TS active site (70). Using computer graphics, naphthalene was visualized and manipulated within this favorable site (Fig. 10.7). This facilitated alterations of the naphthalene scaffold to a benz[*cd*]indole to provide hydrogen-bonding groups to interact with the enzyme and a tightly bound water. Elaboration from the opposite edge of the naphthalene core to extend into the top of the

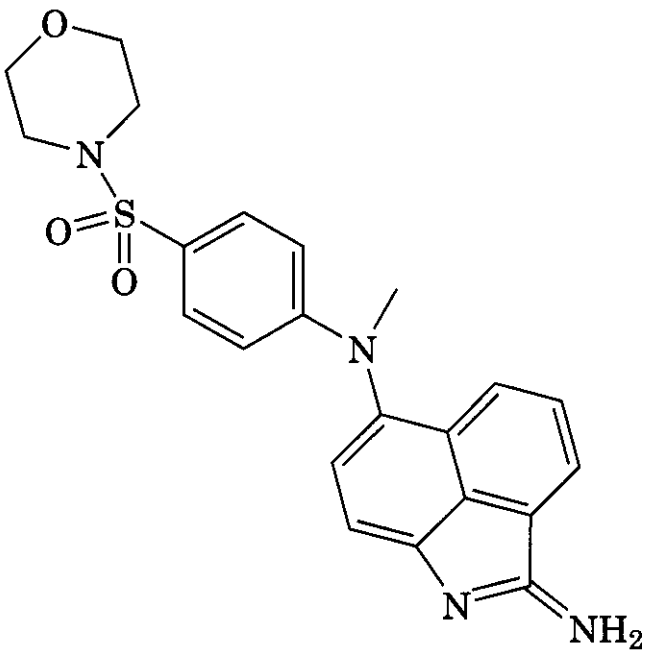
active site cavity, toward bulk solvent, resulted in (20). The use of an amine for the groups attached to position 6 of the benz-



(20)

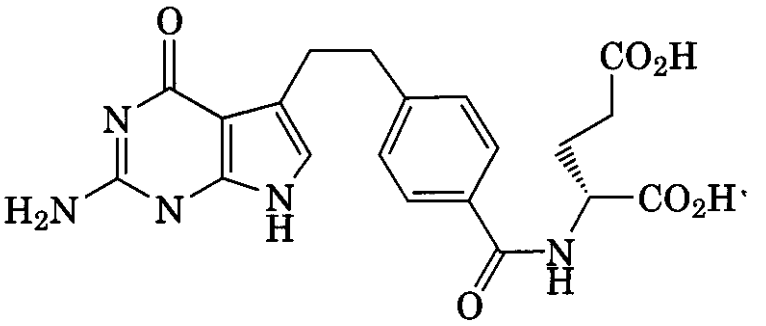
[*cd*]indole improved the synthetic ease for variation of these groups. Compound (20) had a K_{is} value of 3 μM for inhibition of human TS and was about 10-fold less potent against the bacterial enzyme.

The X-ray structure of (20) bound to E. *coli* TS revealed that the compound actually binds more deeply into the active-site crevice than had been anticipated. Instead of interacting favorably with the enzyme-bound water indicated in Fig. 10.7, the oxygen at position 2 of the benz[*cd*]indole displaces it. This forced the Ala263 carbonyl oxygen to move by about 1 Å. Replacement of the oxygen at position 2 with nitrogen provided a significant increase in inhibitory potency. Structural studies revealed that this also resulted in recovery of the displaced water, and restoration of the original position of the Ala263 carbonyl oxygen. The substituents at position 5, on the tertiary amine nitrogen, and on the sulfonyl group were also varied during the iterative optimization process. The process yielded (21) (AG331), which has a K_{is} value of 12 nM for inhibition of human TS. Compound (21) entered clinical trials as an antitumor agent (71).

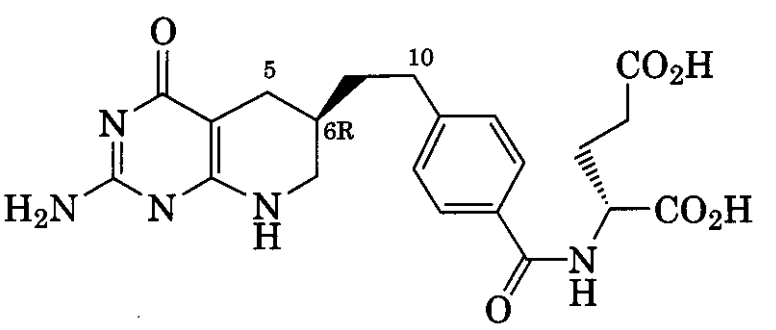


(21)

2.3.3 Glycinamide Ribonucleotide Formyltransferase. Glycinamide ribonucleotide formyltransferase (GARFT) catalyzes the N-formylation of glycinamide ribonucleotide, through use of *N*-10-formyltetrahydrofolate as the one-carbon donor. Because this is an essential step in the synthesis of purine nucleotides, GARFT is a target for blocking the proliferation of malignant cells. Several potent GARFT inhibitors, such as pemetrexed (22, ALIMTA, LY231514) and lometrexol (23, 5,10-dideaztetrahydrofolate, LY-264618), have been shown to be effective antitumor agents in clinical trials (71, 72).



(22) pemetrexed



(23) lometrexol

These were designed through traditional medicinal chemistry approaches, in which an-

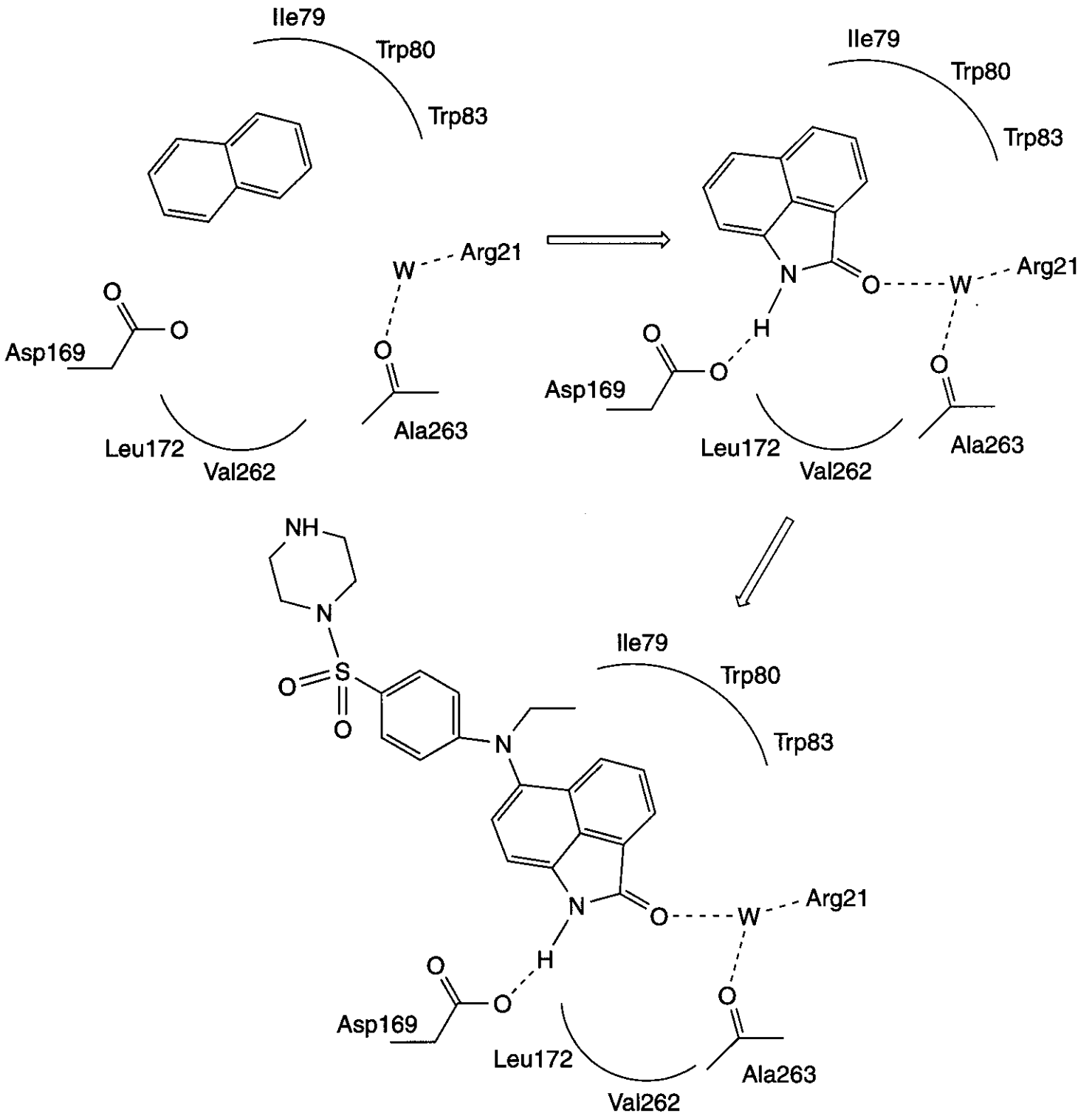
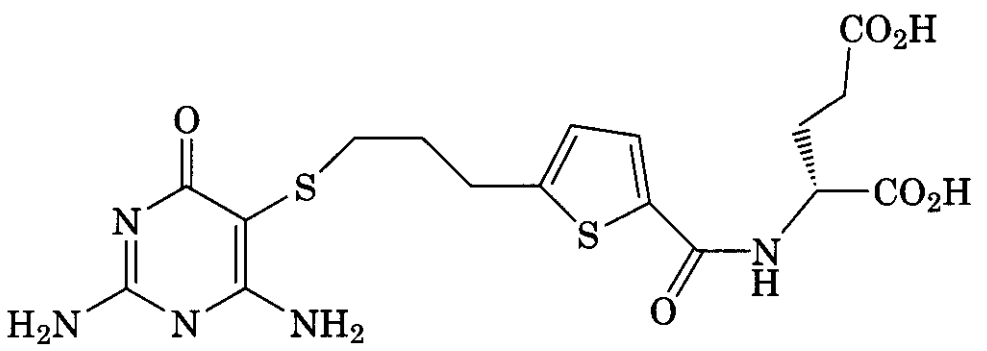


Figure **10.7.** Conceptual design of compound (**20**), by use of the active site of E. *coli* TS as a template. W represents a tightly bound water molecule. [Adapted from Babine and Bender (9).]

alogs of folate were synthesized and then tested as inhibitors of tumor cell growth or of the activity of various folate-dependent enzymes (73–75). A recent paper reported the formation in *situ* of a potent bisubstrate analog inhibitor of GARFT, from glycinamide ribonucleotide and a folate analog, apparently catalyzed by the enzyme itself (76). The substrate analog was designed based on consideration of enzyme structure and the GARFT mechanism. This emphasizes the potential to exploit the interplay between binding and catalytic events in the design of new inhibitors.

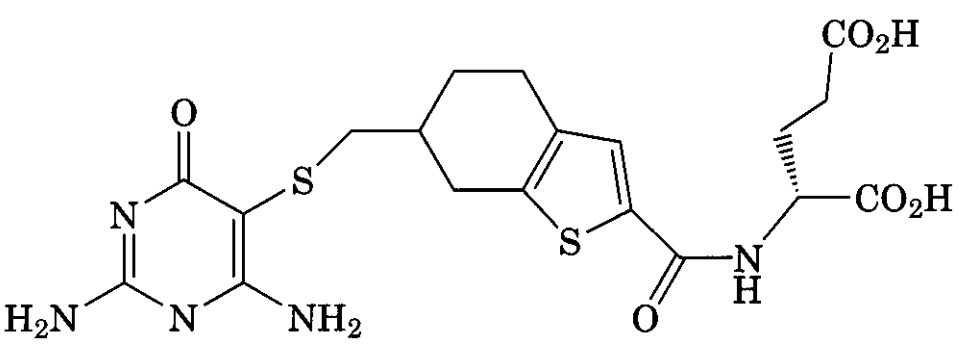
The development of GARFT inhibitors at Agouron began with consideration of the structure of the complex between the E. *coli* enzyme and 5-deazatetrahydrofolate (77). An active and soluble fragment of a multifunctional human protein that contained the GARFT activity was provided by recombinant approaches (78), and its structure was also solved (79) in complex with novel inhibitors. Comparison of the two structures subsequently validated the use of the bacterial enzyme as a model for the human GARFT. The design of novel inhibitors also relied on previous studies of the structure–activity relationships (SAR) for substitutents around the core



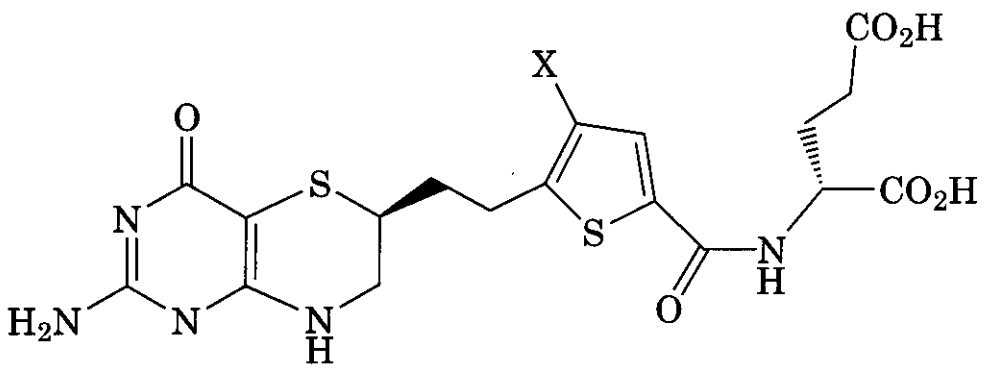
(24)

of (23), including some GARFT inhibitors in which the ring containing N5 was opened (80). Inspection of the structure of the bacterial GARFT–inhibitor complex revealed several important features. The pyrimidine portion of the pteridine was fully buried within the GARFT active site, forming many hydrogen bonds with conserved enzymic groups. The D-glutamate moiety was largely solvent exposed, with no immediately obvious potential for building additional interactions. Retention of the D-glutamate unmodified was also desirable for pharmacodynamic reasons. A significant opportunity was presented by the fact that the active site might accommodate a bulkier hydrophobic atom than the methylene group in 5-deazatetrahydrofolate that replaces the naturally occurring N5 in tetrahydrofolate. To test this idea, a series of 5-thiapyrimidinones were synthesized, including compound (24). These analogs were more readily prepared than the corresponding cyclic derivatives. This compound had a potency of 30–40 n*M* in both a cell-based antiproliferation assay and a biochemical assay for human GARFT inhibition. A crystal structure of human GARFT, complexed with (24) and glycinamide ribonucleotide, confirmed the structural homology between E. *coli* and human enzymes.

Compounds with one fewer methylene in the linker connecting the thiophenyl moiety to the 5-thia position were much less active. Several other analogs, such as (25), were made in attempts to fill the active site more fully, and to restrict the conformational flexibility of the linker. Molecular mechanics calculations failed to correctly predict the conformation on the 5-thiamethylene group of (25) bound to GARFT because of unforeseen conformational flexibility of the enzyme revealed by an X-ray structure of this complex. This again emphasizes the importance of interative experimental confirmation of molecular designs. Several functional criteria in addition to GARFT inhibition and cell-based assays were evaluated during the several cycles of optimization. These included the ability of exogenous purine to rescue cells (which indicates selective GARFT inhibition), and the ability of the inhibitors to function as substrates for enzymes involved in the transport and cellular accumulation of antifolate drugs. Balancing these criteria has resulted in the choice of compounds (26) and (27) (AG2034 and AG2037, respectively) for clinical development at Pfizer. (In 1999, Agouron Pharmaceuticals was acquired by Warner–Lambert, which was subsequently acquired by Pfizer.) It is as yet unclear whether the considerable toxicity of these and other GARFT inhibitors will allow these compounds to be acceptable as anticancer drugs.



(25)



(26) X = H
(27) X = methyl

2.4 Proteases

2.4.1 Angiotensin-Converting Enzyme and the Discovery of Captopril. The design of captopril was a landmark in the application of structural models for developing enzyme inhibitors (81,82). This discovery rapidly led to the development of a family of therapeutically useful inhibitors of angiotensin-converting enzyme for the treatment of hypertension (83). The story has been reviewed thoroughly (for a historical perspective, see either Ref. 84 or Ref. 85), and is briefly summarized here. Angiotensin II, a circulating peptide with potent vasoconstriction activity, is generated by the C-terminal hydrolytic cleavage of a dipeptide from angiotensin I, catalyzed by angiotensin-converting enzyme. Therefore, inhibitors of angiotensin-converting enzyme are vasodilators. ***[An important aside:*** Angiotensin I is generated from a precursor by the action of renin, another exopeptidase that is an aspartyl protease. An orally available renin inhibitor remains an elusive goal, although there are still efforts under way that use SBDD methods (86). Renin inhibitors were early tools in the study of the essential aspartyl protease of human immunodeficiency virus (HIV), which is discussed later.]

Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu → Asp-Arg-Val-Tyr-Ile-His-Pro-Phe + His-Leu
Angiotensin I — Angiotensin II

A key tool in the discovery of captopril at Squibb was the use of a model for the active site of angiotensin-converting enzyme (Fig. 10.8). This model was based on the already known X-ray structure of bovine pancreatic carboxypeptidase A. Both enzymes are C-terminal exopeptidases that require zinc ion for activity, but differ in that carboxypeptidase A releases an amino acid, rather than a dipeptide. Hence, the binding site for the angiotensin-converting enzyme was postulated to be longer, and to contain groups to interact with the central peptide linkage. The suggestion had been made (87) that the inhibition of carboxypeptidase A by benzylsuccinate could be explained by viewing benzylsuccinate as a "by-product analog" (Fig. 10.8, top). The hypothesis was that one of the carboxylates bound into a cationic site, whereas the other interacted with the active site zinc. If this were true, then a similar model for angiotensin-converting enzyme predicted that slightly longer diacids, designed with some regard for the sequence preferences of the converting enzyme, should inhibit that enzyme. This hypothesis was quickly confirmed by the inhibitory activity of succinyl-proline (**28a**).

Peptide sequences related to those of snake venom peptides had already been used to define the structural requirements for peptide inhibitors of angiotensin-converting enzyme. Peptides are unstable *in vivo* and poorly absorbed intestinally, and thus are not good drug candidates. However, the best peptide inhibitor was 500-fold more potent than (**28a**). The

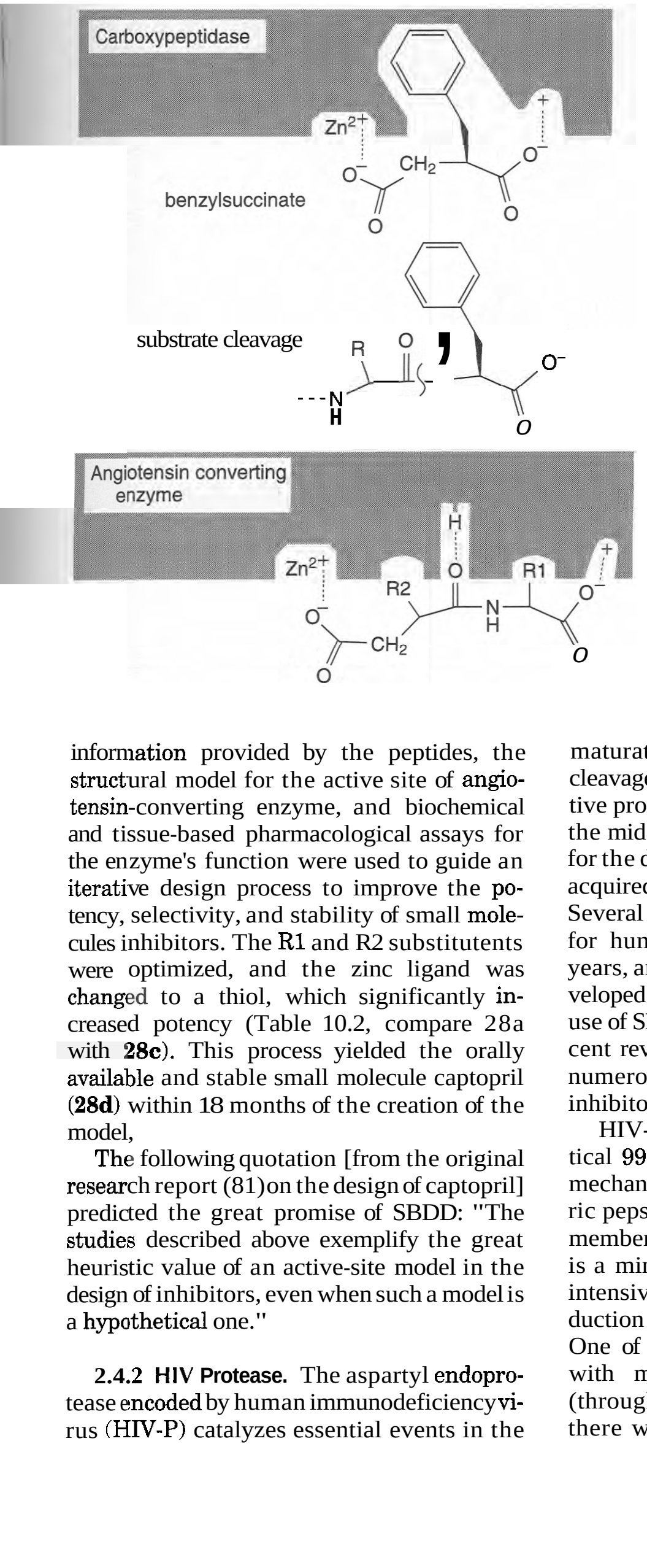


Figure 10.8. Active site models for carboxypeptidase A (top) and angiotensin-converting enzyme (bottom). The design of the dipeptidyl derivative that led to the discovery of captopril is shown bound to the latter enzyme.

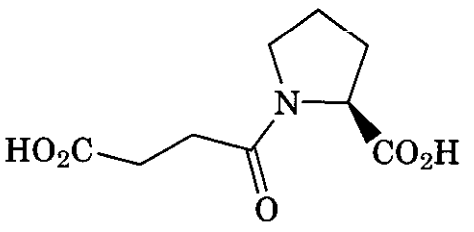
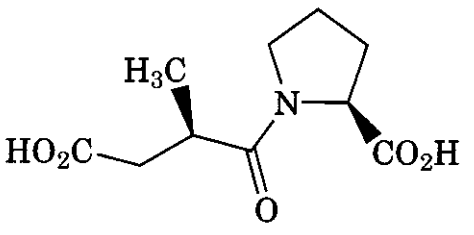
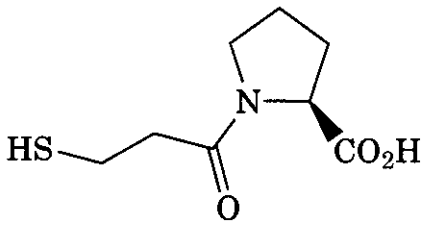
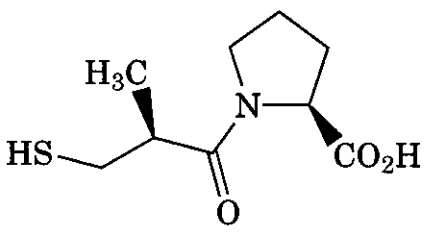
information provided by the peptides, the structural model for the active site of angiotensin-converting enzyme, and biochemical and tissue-based pharmacological assays for the enzyme's function were used to guide an iterative design process to improve the potency, selectivity, and stability of small molecules inhibitors. The R1 and R2 substitutents were optimized, and the zinc ligand was changed to a thiol, which significantly increased potency (Table 10.2, compare 28a with **28c**). This process yielded the orally available and stable small molecule captopril (**28d**) within 18 months of the creation of the model,

The following quotation [from the original research report (81) on the design of captopril] predicted the great promise of SBDD: "The studies described above exemplify the great heuristic value of an active-site model in the design of inhibitors, even when such a model is a hypothetical one."

2.4.2 HIV Protease. The aspartyl endoprotease encoded by human immunodeficiency virus (HIV-P) catalyzes essential events in the maturation of infective virus particles, the cleavage of polyprotein precursors to yield active products. After this was demonstrated in the mid to late 1980s, HIV-P became a target for the development of antiviral drugs to treat acquired immunodeficiency syndrome (AIDS). Several HIV-P inhibitors have been approved for human therapeutic use in the past 10 years, and the speed with which they were developed is attributed in part to the successful use of SBDD methods. There are excellent recent reviews of this area (88, 89). There are numerous reviews of the early work on HIV-P inhibitors (8, 9, 90, 91).

HIV-P is a symmetrical homodimer of identical 99 residue monomers, structurally and mechanistically similar to the pseudosymmetric pepsin family of proteases (92–94), whose members include renin. Because the protease is a minor component of the virion particle, intensive structural studies required overproduction through recombinant DNA methods. One of the first structures was determined with material synthesized nonbiologically (through peptide synthesis). As of June 2002, there were over 100 X-ray structures repre-

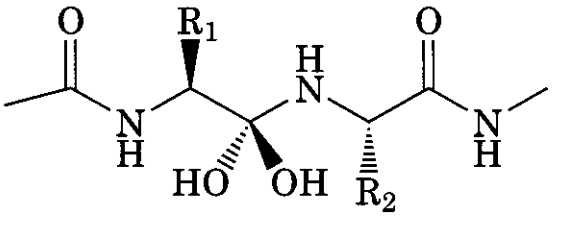
Table 10.2 Key Compounds in the Development of Captopril

Compound	Structure	IC_{50} for inhibition of ACE (μM)
(**28a**) (succinyl-proline)		**330**
(**28b**)		**22**
(**28c**)		**0.2**
(**28d**) (captopril)		**0.023**

sented by coordinate sets in the Protein Data Bank, and many hundreds more have been determined in proprietary industrial studies.

The active site of the enzyme is C2 symmetric in the absence of substrates or inhibitors (Fig. 10.9a), and contains two essential aspartic acid residues (Asp25 and Asp25′). The entrance to the active site is partly occluded by "flaps" constructed of two beta strands (residues 43–49 and 52–58) from each monomer, connected by a turn. In the absence of substrate or inhibitor, the flaps seem to be rather flexible. Upon binding of inhibitors and presumably of substrates, the residues within the flaps undergo movements up to several angstroms to interact with the bound ligand (Fig. 10.10). A single tightly bound water is observed in the structures of most HIV-P–inhibitor complexes, accepting hydrogen bonds from the backbone amides of both flap residues Ile50 and Ile50′ and donating to carbonyls of the bound inhibitors. This is referred to as the "flap" water. Despite the presence of this water and several tightly bound water molecules on the floor of the active site, the cavity also contains extensive hydrophobic surface area. The minor differences between the HIV proteases from two major strains of HIV (HIV-1 and HIV-2) are not addressed here. More significant are the HIV-P sequence variants with much reduced sensitivity to existing drugs that have evolved because of selective pressure and the rapid mutation rate of the virus. The reader interested in the differences between the proteases from HIV-1 and HIV-2, or in the issues surrounding drug-resistant variants, is referred to Ref. 91 and Ref. 89, respectively.

The early work on inhibition of HIV-P was much influenced by previous structural and mechanistic work on pepsin and its inhibitors. Both enzymes are thought to catalyze peptide hydrolysis through a tetrahedral transition state, shown below as (29). The previous work



(**29**)

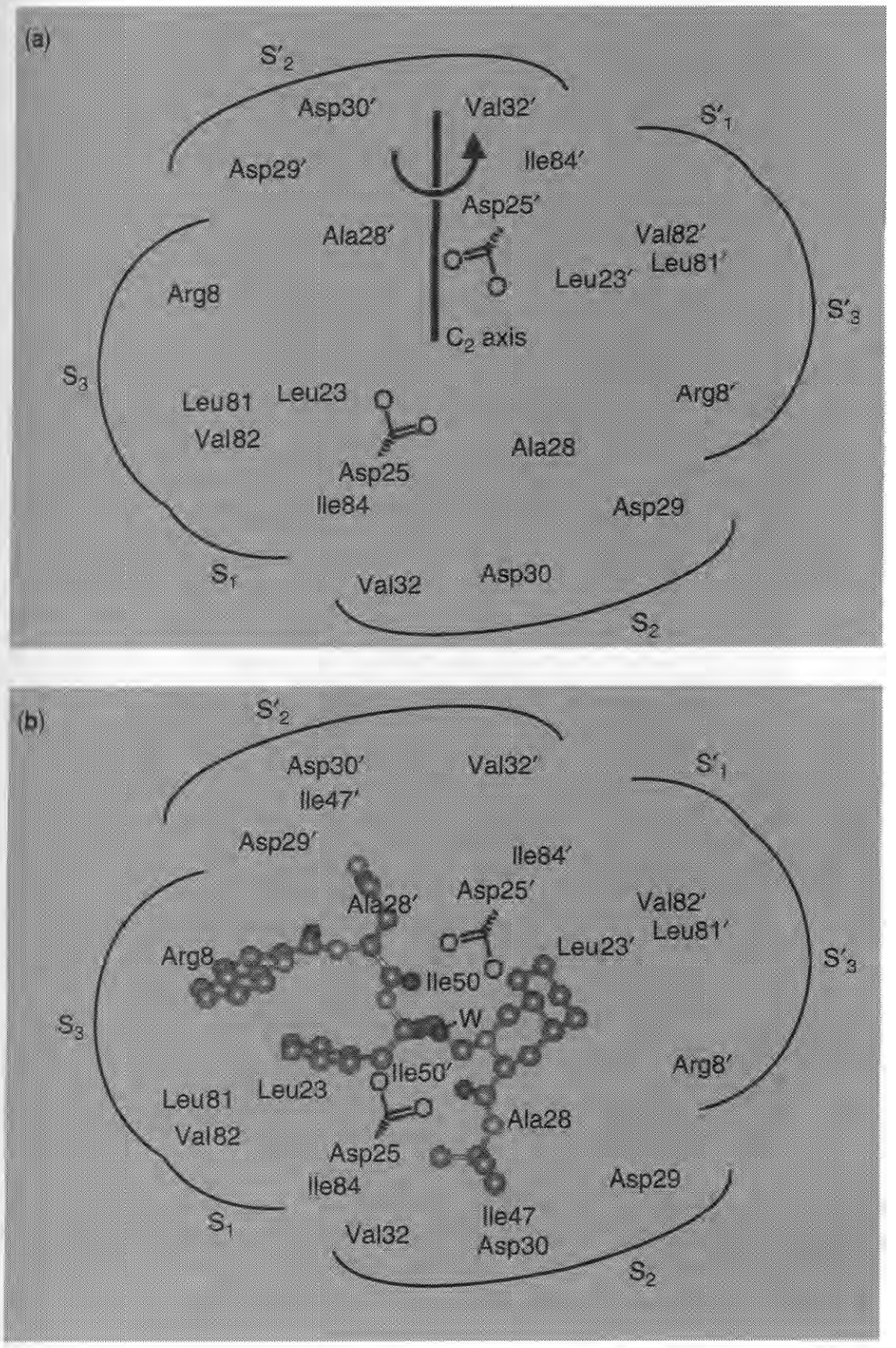
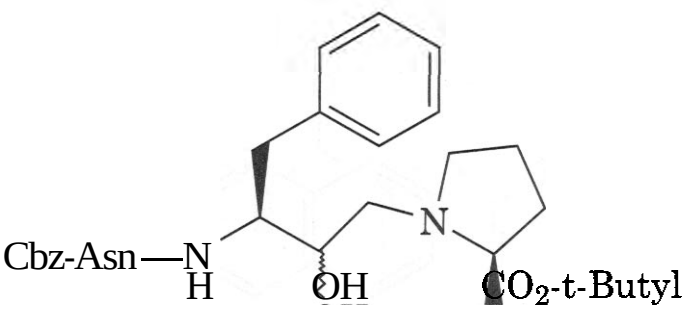


Figure 10.9. (a) Residues in the active site of H N protease. The C2 axis that relates the residues of the two monomers is indicated. The carboxylates of Asp25 and Asp25′ are the catalytic groups. Not shown in this view are several flap residues (Ile47/Ile47′, Ile50/Ile50′), which move in to interact with inhibitors. (b) Active site with bound (**31**) [saquinavir (PDB code 1HXB)]. Note the asymmetry of inhibitor binding. The flap water that is shown very close to saquinavir is labeled W. See color insert.

on transition state mimics as pepsin inhibitors and the sequence of some cleavage sites for HIV-P led to the discovery at Roche of the R and S versions of (**30**) as submicromolar inhibitors of HIV-P, with the R enantiomer being threefold more potent (95). These inhibitors employ a hydroxyethylamine moiety to replace the P1–P1′ linkage that is normally cleaved (the scissile bond) with a stable group. The lead molecules were optimized without knowledge of the HIV-P crystal structure, to produce (31)(Ro 31–8959, saquinavir, Fortovase).



(30)

Saquinavir (31)was the first HIV-P inhibitor approved for human use. Figure 10.9B

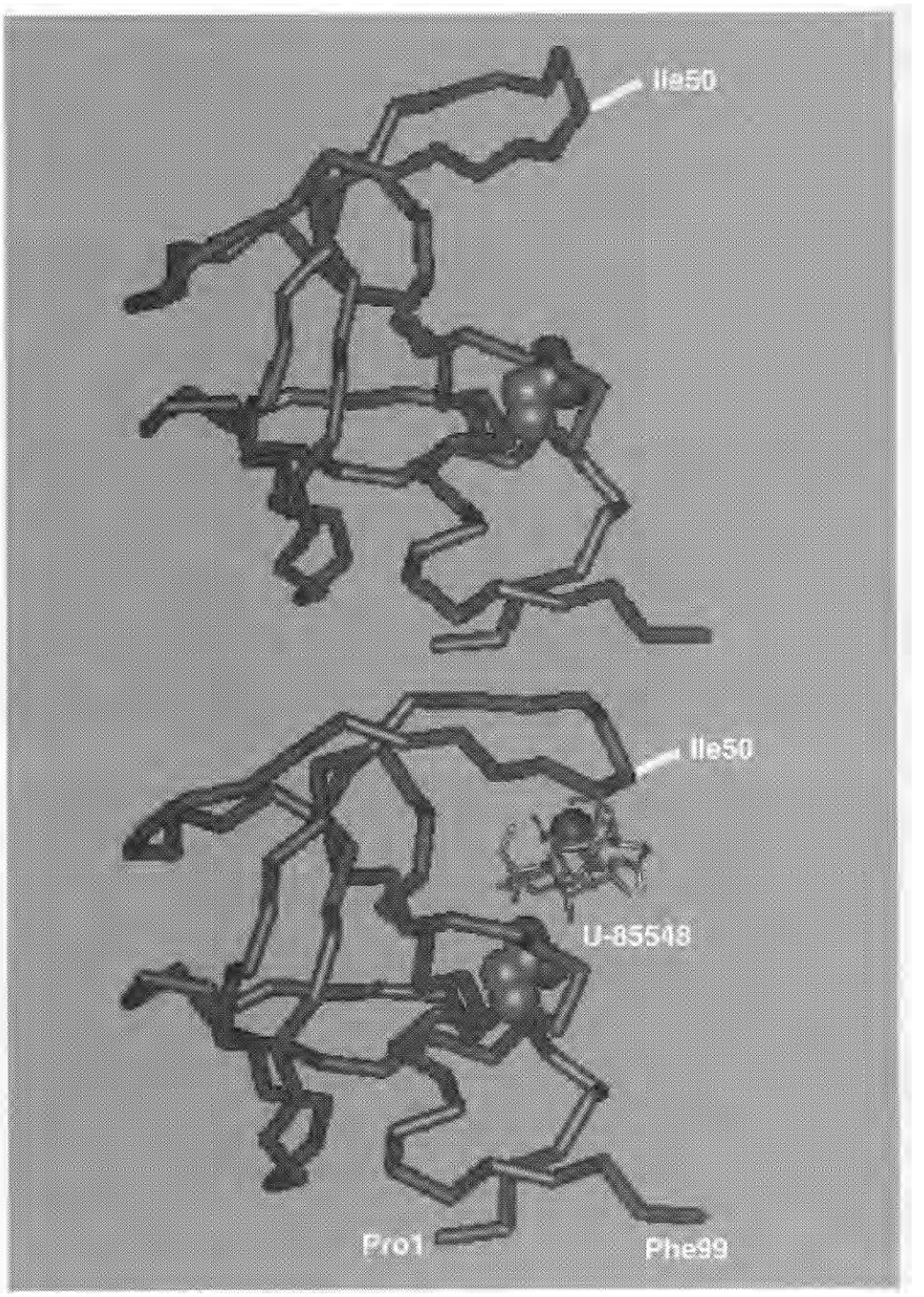
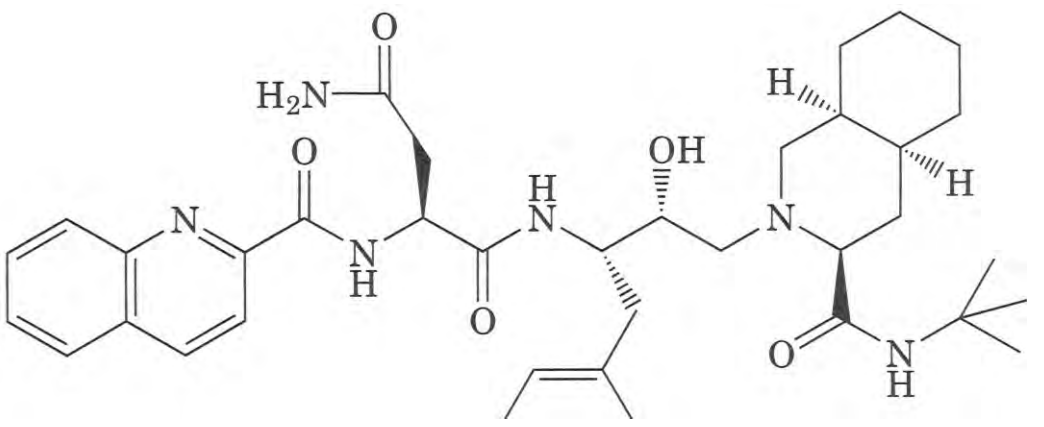
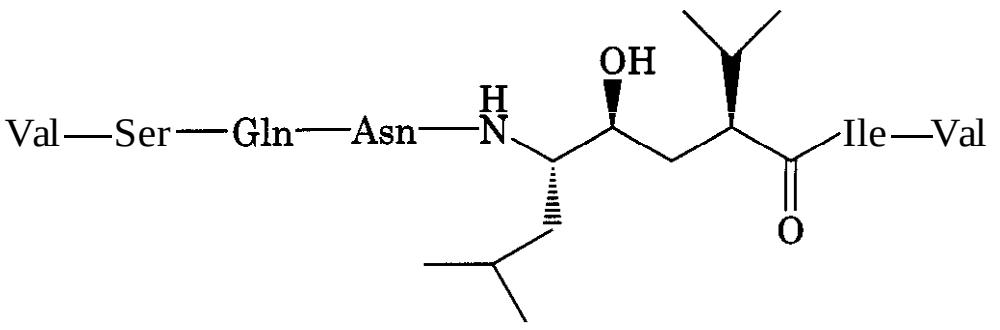


Figure 10.10. Comparison of the structures of HIV-P apoenzyme monomer (top, PDB code 3PHV) and the complex between HIV-P and **(32)** (U-85548; bottom, PDB code 8HVP). The inhibitor is shown as a ball and stick structure. Note the rearrangement of the flap residues; Ile50 is indicated for reference. The van der Waals surface of Asp25 is shown in both structures. The flap water (red ball) is also shown between Ile50 and U-85548. In the bottom structure, the locations of the *N* and C termini of HIV-P are noted. See color insert.

shows the asymmetrical binding mode of the molecule in the HIV-P active site. Because the metabolic and pharmacokinetic characteristics of this compound and several other early HIV-P inhibitor drugs are less than ideal, the search for better ones has continued. Many of the deficits arise from the large size and peptidic nature of the inhibitors. Another early



(31) saquinavir, Ro 31-8959



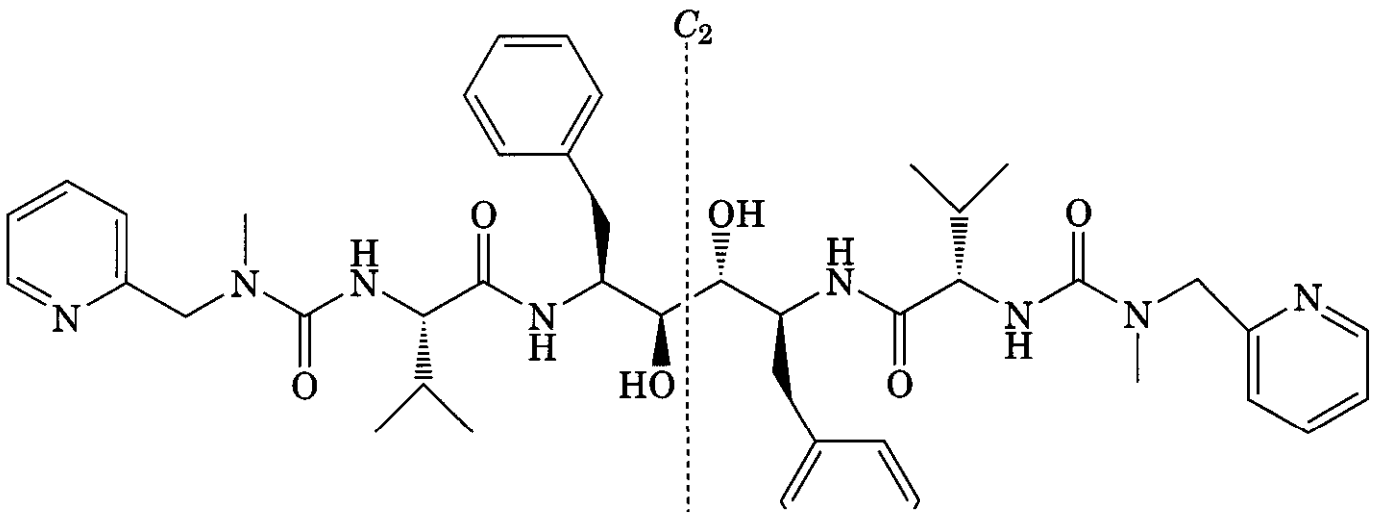
(32)

inhibitor was the modified octapeptide (32, U-85548) developed at Upjohn (96).

This subnanomolar inhibitor was used to define the extensive hydrophobic and hydrogen bonding interactions available in the HIV-P active site (97). A common feature in the binding of (31)and (32) to HIV-P is the interaction of the central hydroxyl group of the inhibitors with the carboxylates of both Asp25 and Asp25′. This hydroxyl group replaces a water molecule that likely binds between these aspartyl side chains during peptide hydrolysis by HIV-P. The inhibitors can therefore be seen as mimics of a "collected substrate." The liberation of this water to bulk solvent probably contributes about 5 kcal mol^{-1} to the free energy of inhibitor binding, based on the studies by Rich and his colleagues on similar inhibitors of pepsin (98, 99). An interesting difference between (31)and (32) is that (31) has R stereochemistry at the hydroxymethyl center, whereas in (32) this is an S center. Part of the reason for this is that when (31) binds to HIV-P, the decahydroquinoline ring system induces a conformational change in the protein, affecting primary site S,'. The optimal stereochemistry at the hydroxymethyl center appears to be whichever one will allow the interaction of the hydroxyl with both catalytic aspartates while accommodating the placement of inhibitor moieties in the S,, S,, S,', and S,' sites with minimal conformational strain on the inhibitor (9).

Both (31)(Fig. 10.9b) and (32) (Fig. 10.11) bind to the HIV-P active site asymmetrically. However, after the X-ray studies of crystalline HIV-P apoenzyme revealed it to be a symmetrical dimer, C2 symmetric inhibitors were designed to take advantage of this structural feature (Fig. 10.12). Both alcohol diarnines and diol diamines were examined. For example, the C2 symmetric compound (33) (A-77003) was synthesized at Abbott and entered clinical trials as an antiviral agent for intravenous treatment of AIDS (100).

The X-ray structures of complexes between HIV-P and diol diamine derivatives like (33) showed (101) that, although one of the hydroxyl groups bound between the catalytic aspartyl carboxylates and made contacts with both, the second hydroxyl made only one such



(33) A-77003

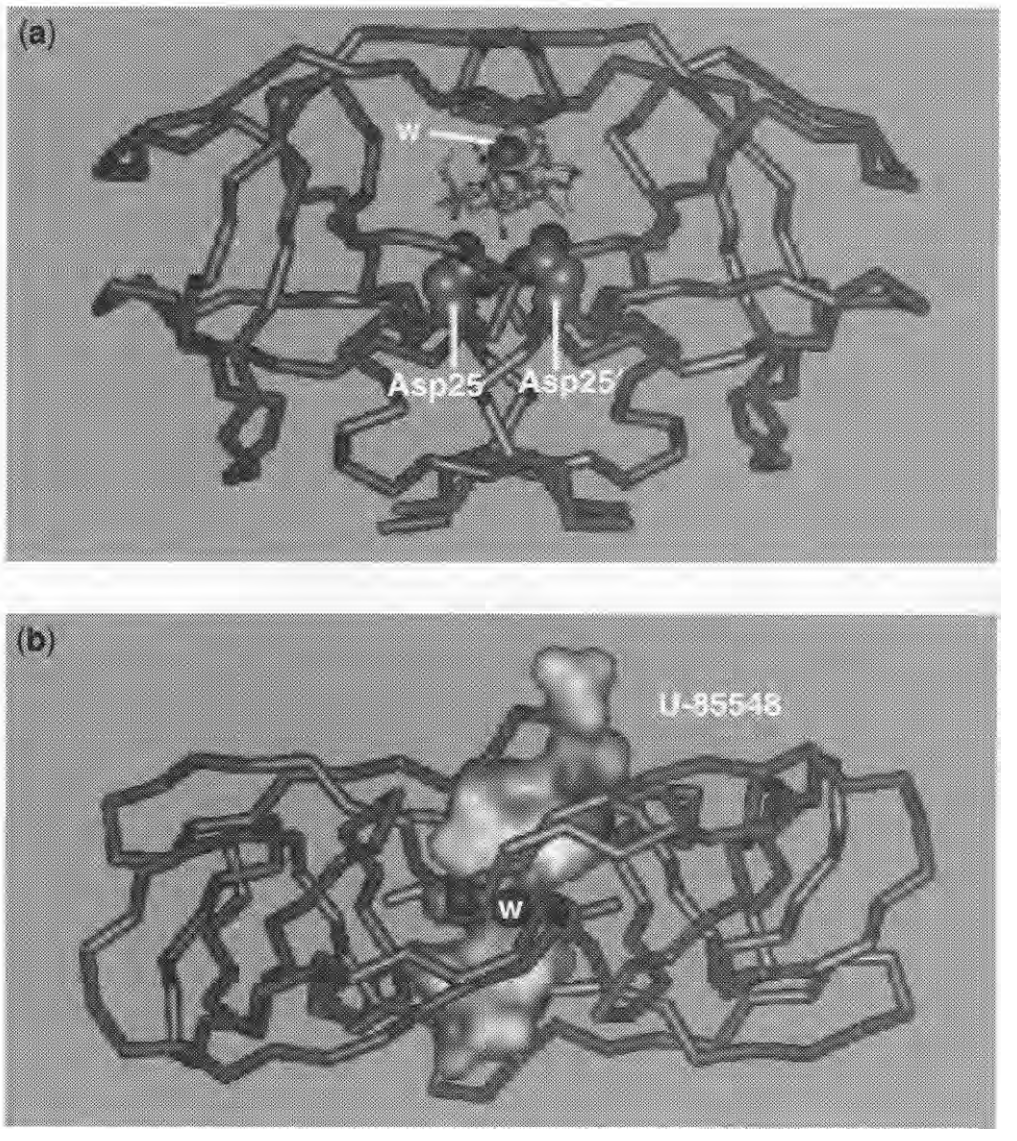


Figure 10.11. Orthogonal views of the complex between HIV-P and **(32)** (U-85548). The view in panel a is rotated approximately 90" (around the long axis of the protein) from the view in panel b. Van der Waals surfaces of Asp25, Asp25′, and the flap water (W) are shown. In panel b, the solvent-accessible surface of the inhibitor is shown. See color insert.

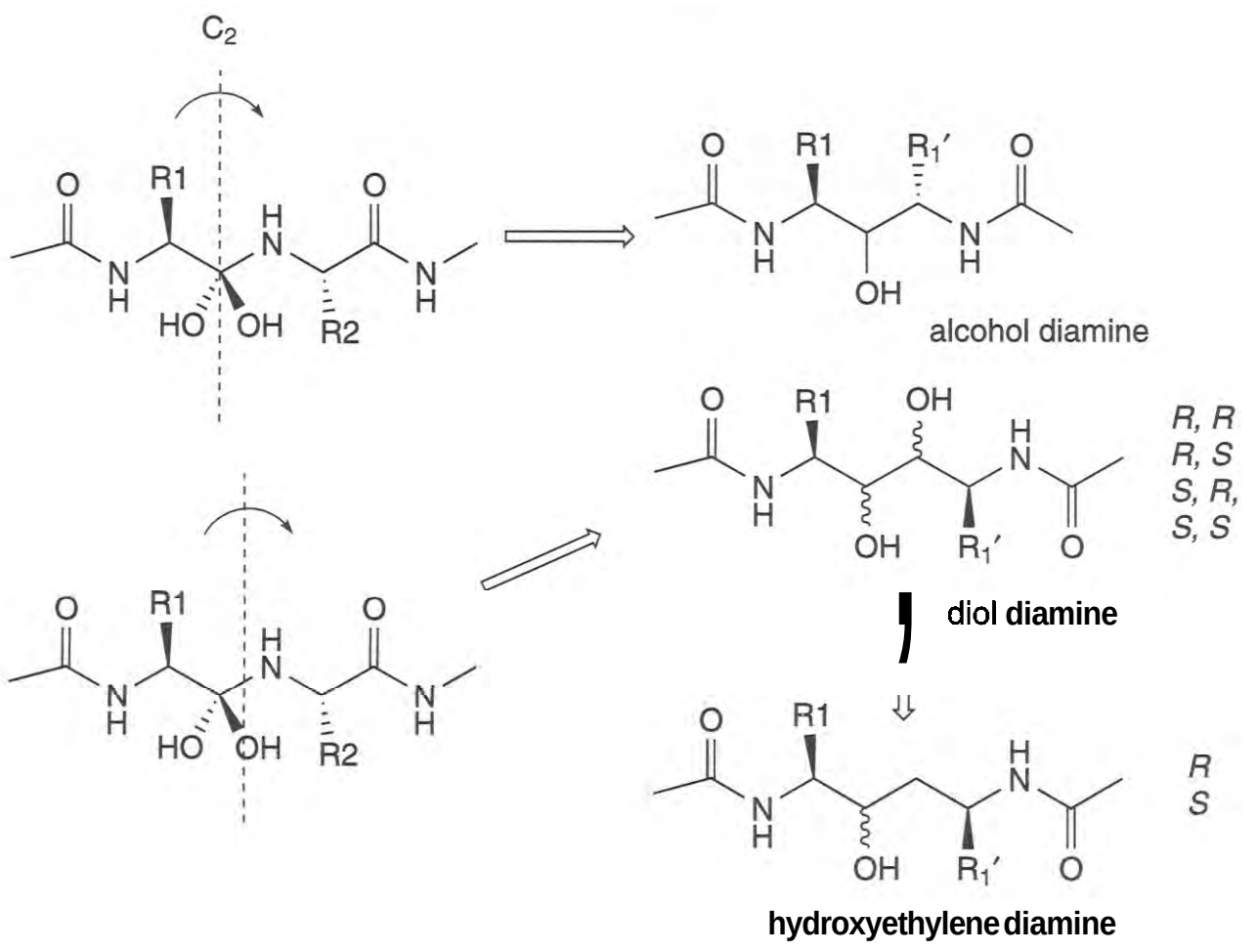
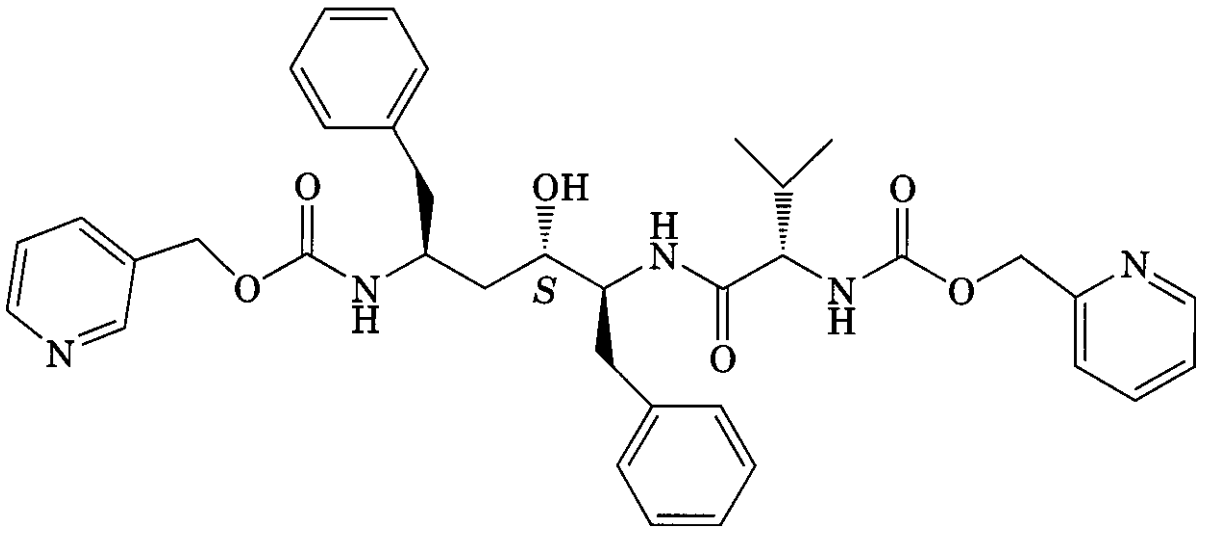


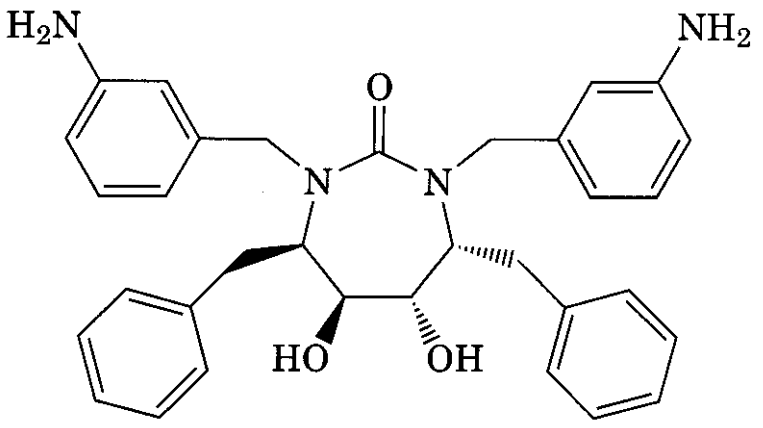
Figure 10.12. Design principle for **C2** symmetric inhibitors of HIV-P and the related hydroxyethylene diamine scaffold.



(34) A-80987

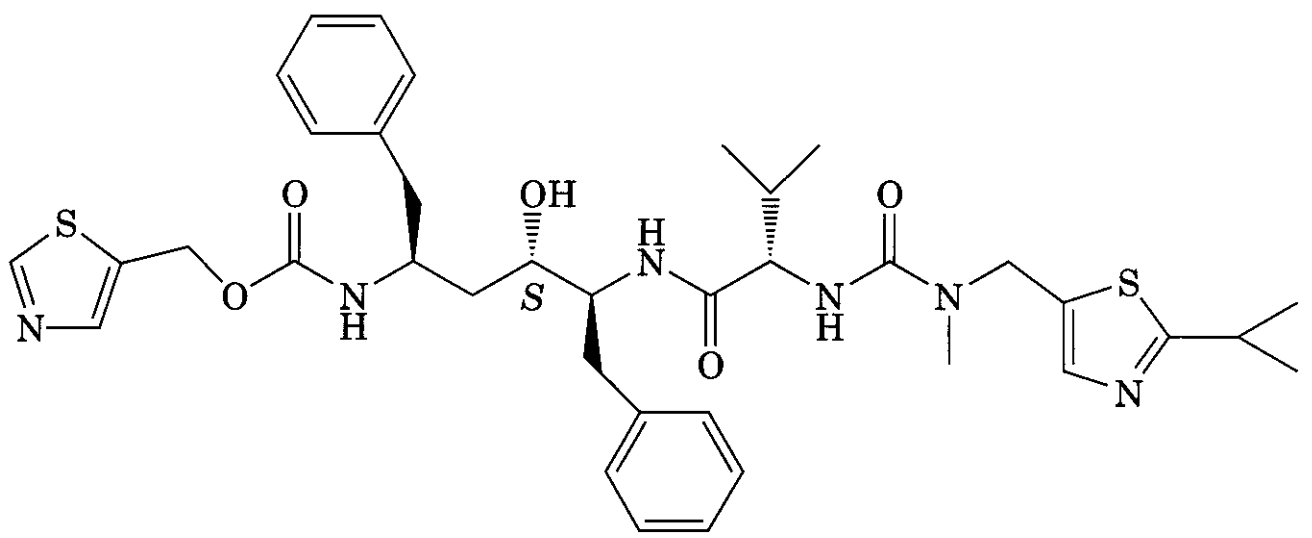
contact. Thus the cost of desolvating the second inhibitor hydroxyl upon binding is not compensated by strongly favorable interactions in the complex (8). This led to the deletion of the second hydroxyl, as seen in compound **(34)**, another compound in this program at Abbott. Further structural modifications, to enhance solubility and metabolic stability, were guided by the fact that the "ends" of the protease-bound inhibitors were relatively solvent exposed and made fewer contacts with the enzyme (102). Deletion of a valine residue (33→34) gave a smaller compound, presumably aiding solubility and absorption. The eventual product of this program was ritonavir (**35**, A-84538, ABT-538, or Norvir), which has been successfully launched.

Another C2 symmetric HIV-P inhibitor, discovered at Dupont Merck is compound (36) (DMP-450). This was one of a series of cyclic ureas designed to interact with both the aspartyl carboxylates and the Ile50 and Ile50′ backbone amides that hydrogen bond with the flap water (103). The compounds interacted with HIV-P in a highly symmetrical fashion, as they had been designed to do, with the urea oxygen replacing the flap water. Compound (36) was licensed to Triangle Pharmaceuticals, and the mesylate advanced into Phase I clinical trials. Its future is uncertain after the trials were put on hold because of animal toxicity (http://www.tripharm.com/dmp45O.html).

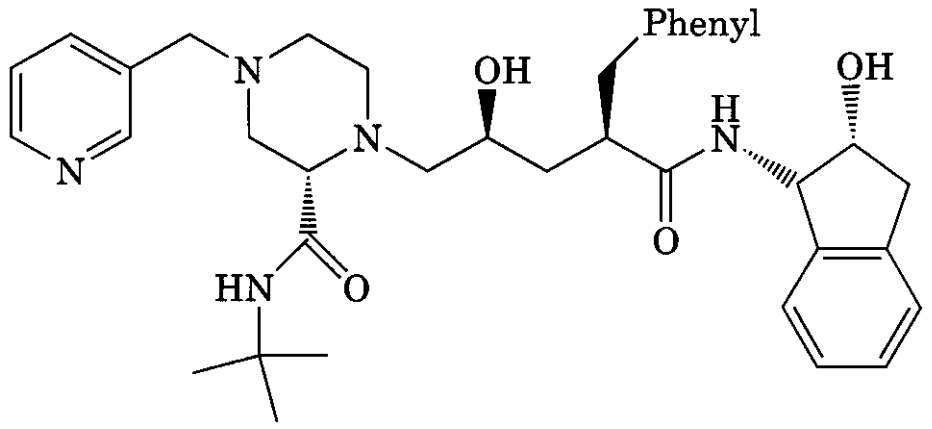


(36)

One of problems common to many of the HIV-P inhibitors already discussed is their



(35) ritonavir



(37) indinavir

low solubility, which translates to low bioavailability. The discovery of (37) (indinavir, L-735,524) was the result of the successful application of SBDD at Merck to directly address this problem. During an iterative optimization process, the physicochemical properties of HIV-P inhibitors were modified within constraints that were established structurally (104). Crixivan (the sulfate of 37) was successfully launched for use as an antiviral drug.

The process leading to indinavir (Fig. 10.13) began with (**38**), a hydroxyethylene-containing heptapeptide mimic, originally designed as a renin inhibitor (105). The inhibition of HIV-P by (38) was discovered by screening. Classical medicinal chemistry methods allowed a reduction in size, and the discovery of an amino-2-hydroxyindan moiety to replace the terminal dipeptide (corresponding to P_2', thought to bind into the S_2' site). This approach (105, 106) resulted in the generation of (39)(L-685,434). Although (39) had a subnanomolar IC_{50} for inhibition of HIV-P, it also had very low aqueous solubility, like most peptidomimetics. One way to improve solubility is to insert a charged functional group into the molecule. The tertiary amino group in the HIV-P inhibitor saquinavir (31)

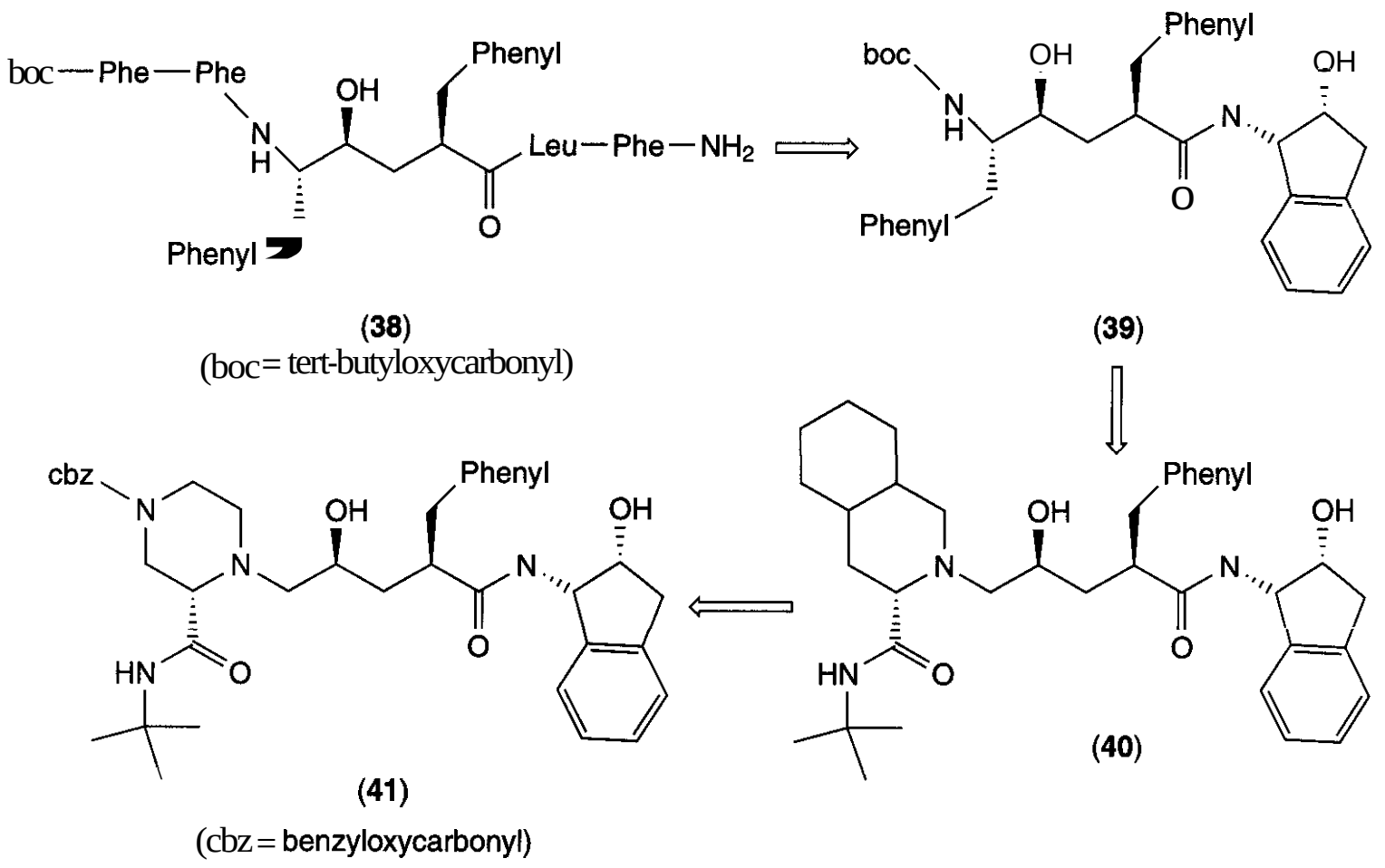
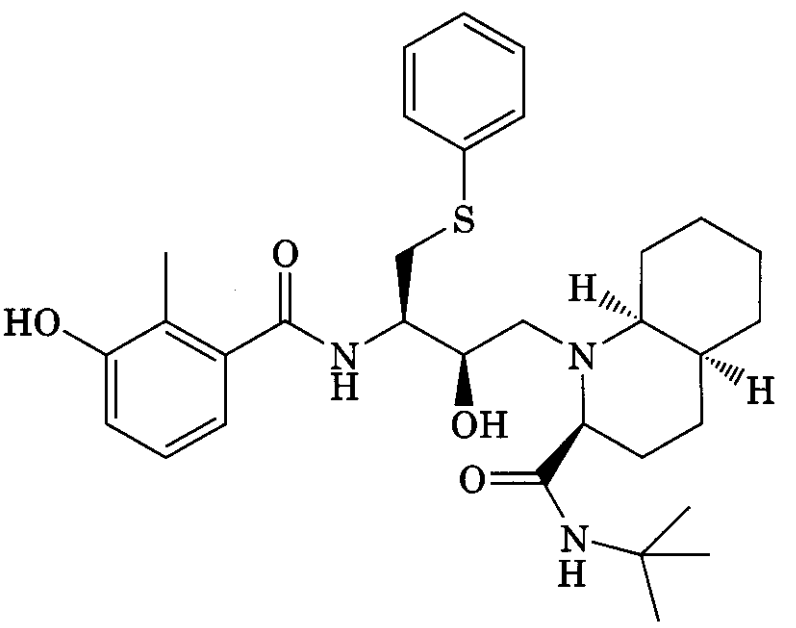


Figure 10.13. Structures of HIV-P protease inhibitors during the optimization process leading to the discovery of (**37**) (indinavir).

was already identified. Piracy of the decahydroisoquinoline tert-butylamide from (31) provided the idea for the hybrid molecule (40). In addition to the charged group, use of this ring system would partly "preorder" the inhibitor's structure, lessening the entropic cost of binding. Molecular modeling was used with known structures of HIV-P–inhibitor complexes to evaluate this idea, and it was judged to be reasonable enough to justify the synthesis of (40) (104). This compound was subsequently shown to have much better pharmacokinetic behavior than its antecedents, consistent with improved solubility and dissolution.

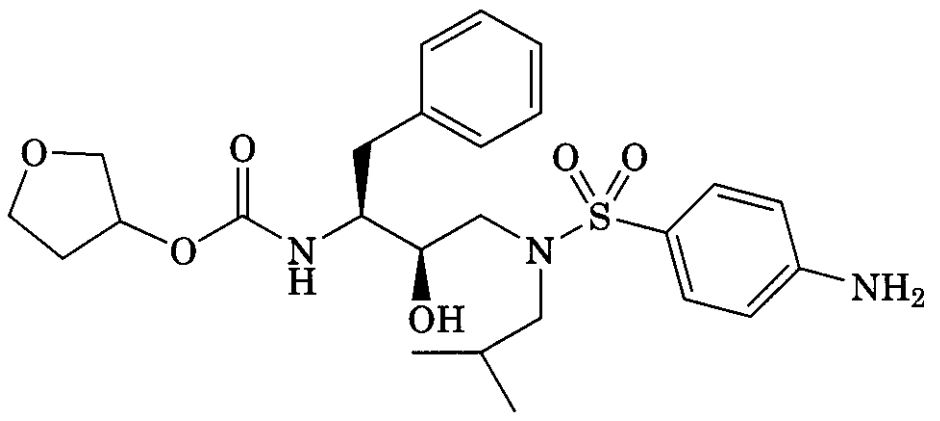
A convergent synthetic route was devised to generate (40) to improve the accessibility of important analogs. Although (40) was an 8 n*M* inhibitor of the isolated enzyme, better potency was needed for acceptable cell-based activity, and still better solubility characteristics were needed. A method for structure-based computational estimation of the interaction energy for HIV protease inhibitors with the enzyme was developed and used to help estimate inhibitor potency before synthesis (107). Variation of the group contributing the tertiary amine led to the discovery of the piperazine derivative (41) (L-732,747), which had subnanomolar potency against HIV-P. The X-ray structure of the HIV-P complex with (41) confirmed the binding mode predicted by molecular modeling, with the molecule filling the S_1, S_1, S_1', and S_2' pockets, and the S_3 pocket occupied by the terminal benzyloxycarbonyl moiety. Replacement of the benzyloxycarbonyl with more polar heterocycles, chosen to be accommodated by the S_3 pocket and to further improve aqueous solubility, yielded (37).

Several other approved AIDS drugs that act by inhibition of HIV-P have also been developed through use of SBDD methods. Compound (42) (amprenavir, Agenerase, also known as VX-478) is the most recent addition to the HIV-P inhibitors approved for human antiviral treatment, and differs significantly from earlier inhibitors. Compound (42) was specifically designed by Vertex scientists to minimize molecular weight to increase oral



(43) nelfinavir

bioavailability (108). Compound (43) (nelfinavir, AG-1343, also known as LY312857), like the precursors to the earlier drug (37) (indinavir), copied the decahydroisoquinoline *tert*-butylamide group from the first marketed HIV-P inhibitor (31) (saquinavir). Compound (43) was developed in a collaboration between scientists at Lilly and Agouron (109), and is mar-



(42) amprenavir

keted by Pfizer as Viracept, the mesylate salt of nelfinavir. In both (42) and (**43**), the scientists involved used iterative SBDD methods to alter the physicochemical properties of the drug molecule while maintaining potency by optimizing interactions with the active site of the enzyme. An important feature shared by these compounds is the fact that the bound inhibitors appear to be in low energy conformers, so that minimal conformational energy costs must be paid upon binding to the enzyme.

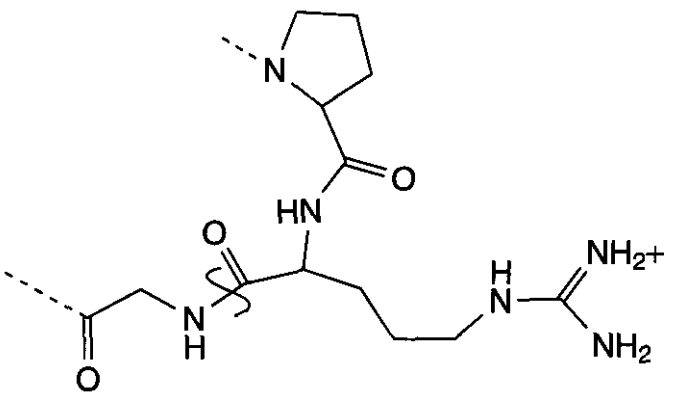
2.4.3 Thrombin. Thromboembolic diseases such as stroke and heart attack are major health problems, especially in many Western countries. This has led to searches for drugs that are effective inhibitors of various serine endoproteases in the blood-clotting cascade, such as factor Xa and thrombin. Existing therapeutic agents such as the coumarins (like warfarin), heparin, and hirudin have problems related to their absorption or unpredictable metabolism and clearance. Recently, new small molecule inhibitors of thrombin have become available for human use in the United States, including (44) (argatroban, MD-805, developed by Mitsubishi) and (45) (melagatran, developed by AstraZeneca) (110, 111). These nanomolar inhibitors of human thrombin were optimized by classical medicinal chemistry, starting with peptidomimeticssimilar to the thrombin cleavage site in fibrinogen (see Fig. 10.14a). Poor absorption by an oral route requires that they must be administered intravenously or at best subcutaneously. At present, the only direct inhibitor of thrombin suitable for oral administration is ximelagatran, a prodrug form of melagatran in late development for various cardiovascular indications by AstraZeneca as of mid-2002. The therapeutic need and the availability of high quality crystal structures for human thrombin bound to inhibitors such as (44) make this an attractive target for SBDD (112). The significant efforts at Merck to use SBDD approaches to develop orally available inhibitors of thrombin, which have yielded compounds that have entered clinical trials, have been reviewed (113,114). For a good overview of this area, see the review by Babine and Bender (9).

Compound (46) [NAPAP, N-alpha-(2-naphthylsulfonylglycyl)-4-amidinophenylalanine piperidide] is a moderately potent inhibitor of human thrombin, but was found to have an unacceptably short plasma half-life in animals (115). However, (46) has been a useful experimental tool and a variety of analogs have been made. The structures of (44) and (46) bound to human thrombin show that they bind somewhat differently, as shown in Figure 10.14b (112,116). However, both form hydrogen bonds with the backbone at Gly216 (part of the oxyanion hole), and both fill the S_1 specificity pocket with a permanent cation attached to an extended hydrophobic group. Compound (46) was the starting point at Boehringer Ingelheim for the development of the orally bioavailable prodrug (47) (BIBR-1048) that generates in *vivo* a potent inhibitor of human thrombin (117). Compound (47) is currently in human clinical trials.

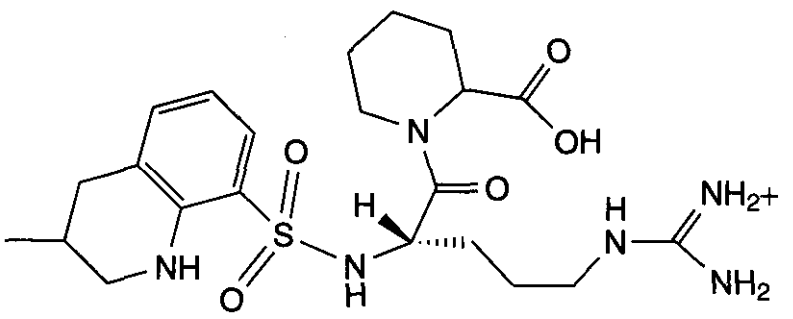
Scientists at Boehringer Ingelheim used the crystal structure of the complex between (**46**) and human thrombin to design a replacement for the central bridging glycine moiety. The hypothesis that a trisubstituted benzimidazole could correctly place groups into the S_1, S_3, and S_4 pockets was confirmed. The first such compound made was (48). The IC_{50} for thrombin inhibition by (48) was only 1.5 μM, but the compound had an improved serum half-life in rats. Determination of the crystal structure of the thrombin–(48) complex showed that (48) binds in a similar fashion to (**46**). The N-methyl on the benzimidazole fit into the P_1 pocket, and the phenylsulfonyl group extended into S_4. The low affinity is likely attributable to the fact that (48) forms no hydrogen bonds with the backbone of Gly216. An iterative optimization process (Fig. 10.15) was used to regain the lost affinity, eventually surpassing the thrombin affinity of the starting point (46) (0.2 μM).

Surprisingly, the N-methyl group could not be replaced with larger alkyl substituents, despite what appeared to be room for them in the P_1 pocket. However, replacing phenyl with larger aryl groups such as naphthyl or quinolinyl on the sulfonamide provided favorable interactions in the P_4 pocket. The crystal structure suggested that the increased lipophilicity of such aryl groups could be balanced by appending charged substituents to

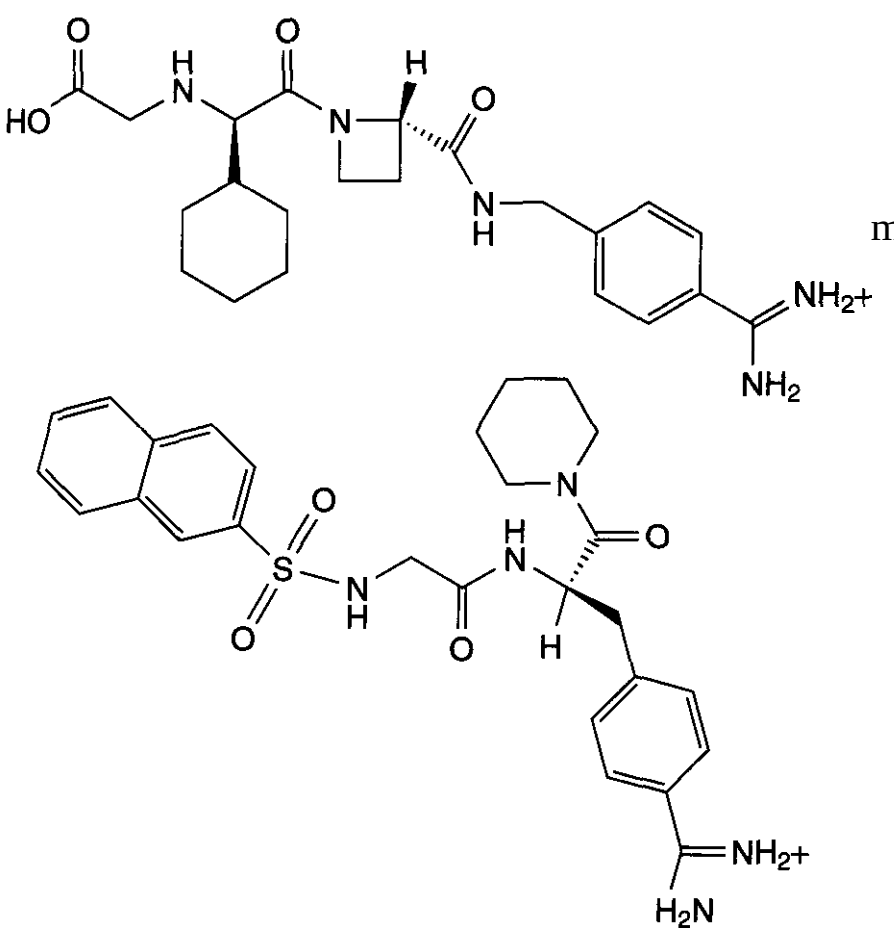
(a)



P_2 P_1 P_1'
-Pro-Arg-Gly-



(44)
argatroban



(45)
melagatran

(46)
NAPAP

Figure 10.14. (a) Sequence in fibrinogen at the thrombin cleavage site (top), and structures of several inhibitors of human thrombin.

the sulfonamide nitrogen. Such substituents appeared likely to extend into solvent and therefore to be tolerated without compromising affinity. This was confirmed (i.e., compound **49**), and this decreased the undesirable affinity for serum-binding proteins. X-ray studies with some of the inhibitors at this point indicated that a longer linker between the central benzimidazole and the benzamidine moiety in the S_1 pocket might provide some advantage. This was confirmed with several analogs, with the methylamino linker providing a 10-fold increase in potency (compound 50). By this point, the structural basis for interaction of this compound series with thrombin was understood sufficiently to suggest that the amidosulfonyl group could be replaced by a carboxamide. This was confirmed by use of several compounds, such as (**51**). Compound (**51**) (BIBR **953**) was quite active as an anticoagulant in animals dosed intravenously, but required conversion to prodrug (compound 47) to mask its charge and allow oral dosing.

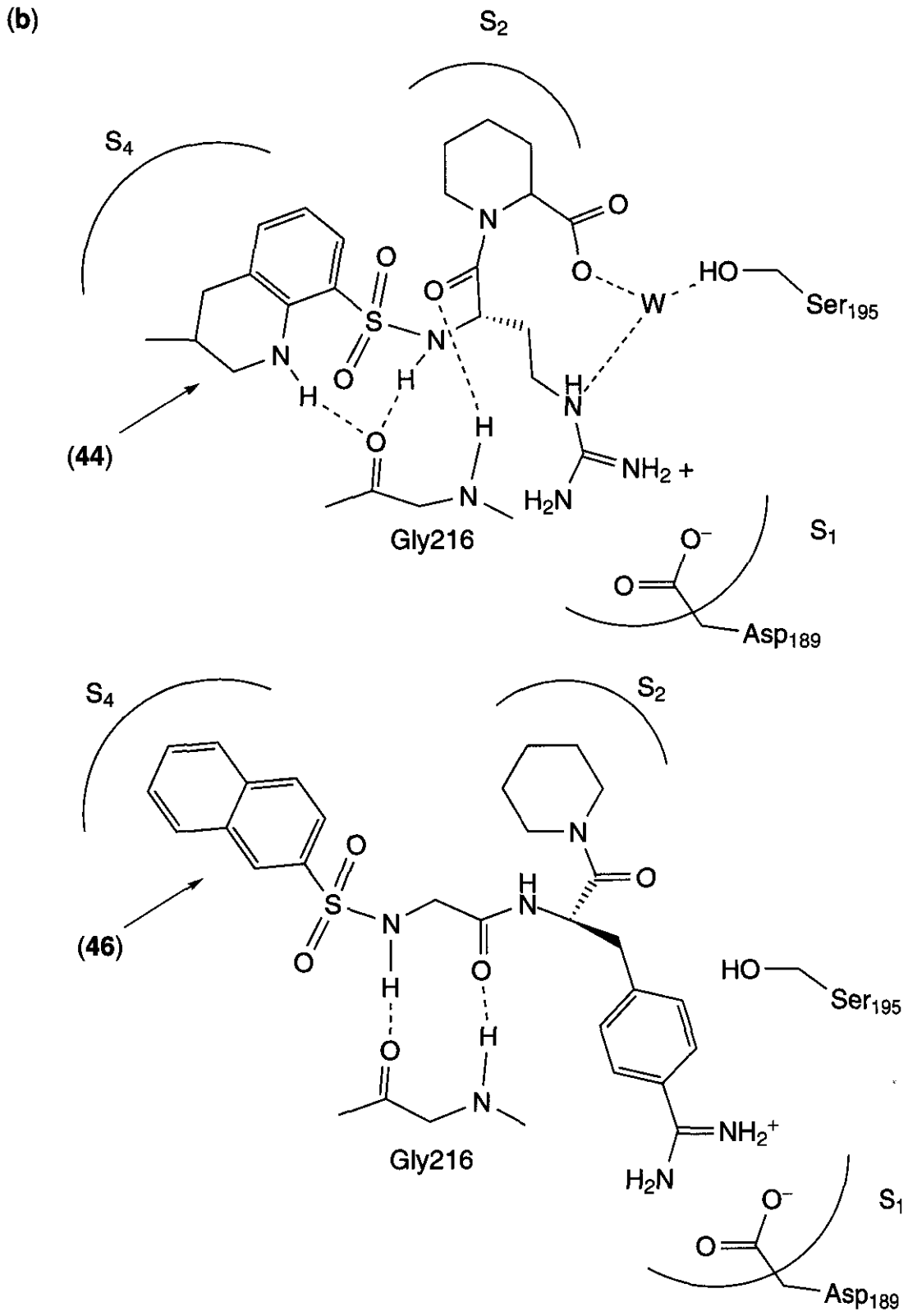
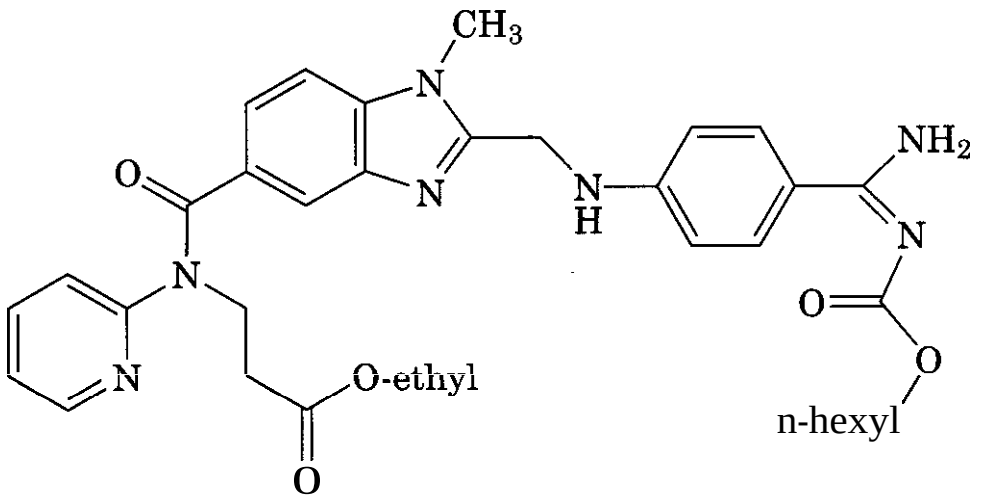


Figure 10.14. (b) Schematic comparison of the binding interactions for (**44**) and (**46**) in X-ray structural models of crystalline thrombin.

2.4.4 Caspase-1. Caspase-1 (interleukin 1-β converting enzyme, or ICE) is a member of a family of cysteine proteases that catalyze the cleavage of key signaling proteins in such processes as inflammatory response and apoptosis. Genetic methods have provided evidence supporting a role for caspase-1 in diseases such as stroke (118) and inflammatory bowel disease (119). The X-ray structure of crystalline human caspase-1 was solved in 1994 by several groups (120, 121), and has been a valuable tool in intensive efforts to design potent and bioavailable inhibitors of the enzyme.

Compound (**52**) (pralnacasan, VX-740) was developed as a caspase-1 inhibitory therapeutic agent through use of SBDD in a collaboration between Vertex and Aventis. Although the details of the discovery process have not been published, (52) probably functions as a prodrug. The cleavage of the lactone of (52) would yield a hemiacetal that could hydrolyze to release ethanol and the aldehyde form of the drug, which then can form a covalent thioacetal with the active site thiol of caspase-1, leading to pseudoirreversible inhibition. Clinical trials of compound (52) as an anti-inflammatory agent for treatment of rheumatoid arthritis began in 1999 (122). In April 2002, the



(47) BIBR **1048**

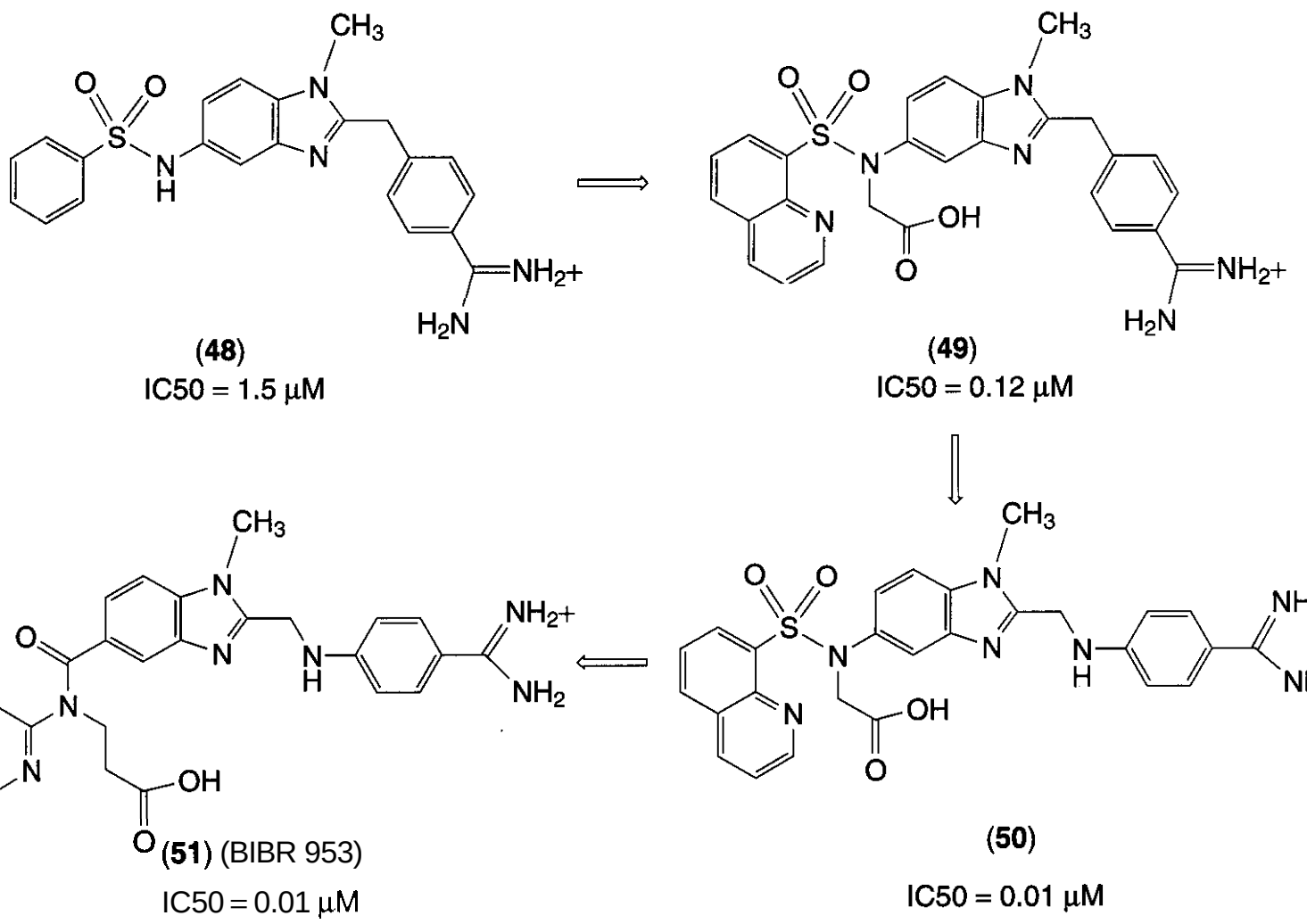
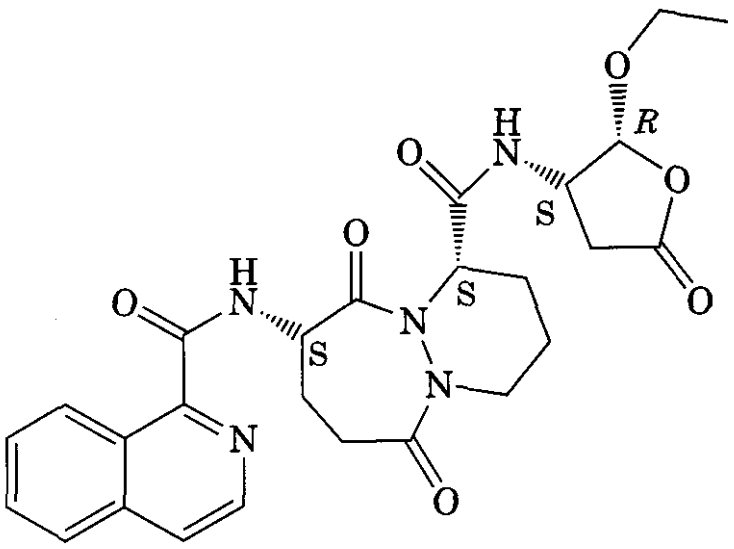


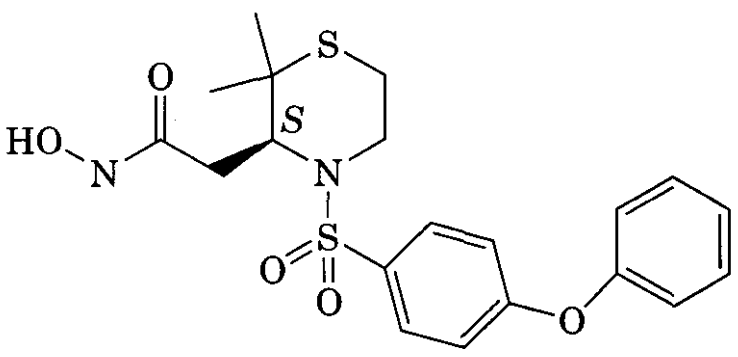
Figure 10.15. Optimization of structure leading to the discovery of **(51)** (BIBR 953).

companies announced that these trials would continue and would be expanded to include treatment of osteoarthritis.

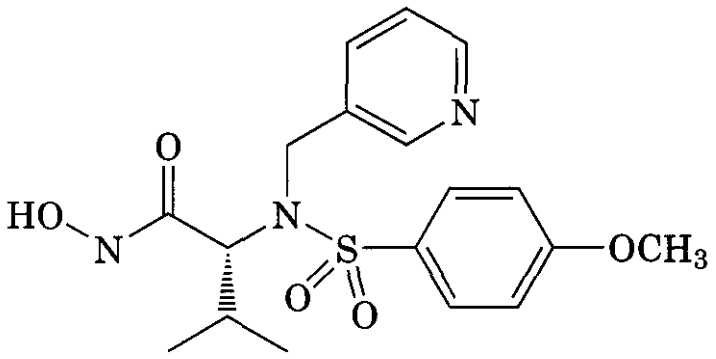
2.4.5 Matrix Metalloproteases. Matrix metalloproteases (MMPs) are a large and diverse family of zinc endoproteases. Several members of this family (such as the collagenases and the stromelysins) are thought to have important roles in proliferative diseases, including arthritis, retinopathy, and metastatic invasiveness of tumor cells. There are publicly available X-ray structures of enzyme–inhibitor complexes for at least seven different MMPs, as of this writing. Several detailed reviews of the SAR and binding modes for inhibitors of matrix metalloproteases are available (9, 123). All MMP inhibitors contain a moiety that binds to the active site zinc, such as the hydroxamates of (53) (prinomastat, AG3340) and (54) (CGS-27023) and the carboxylic acid of (55) (tanomastat, BAY 12–9566). These



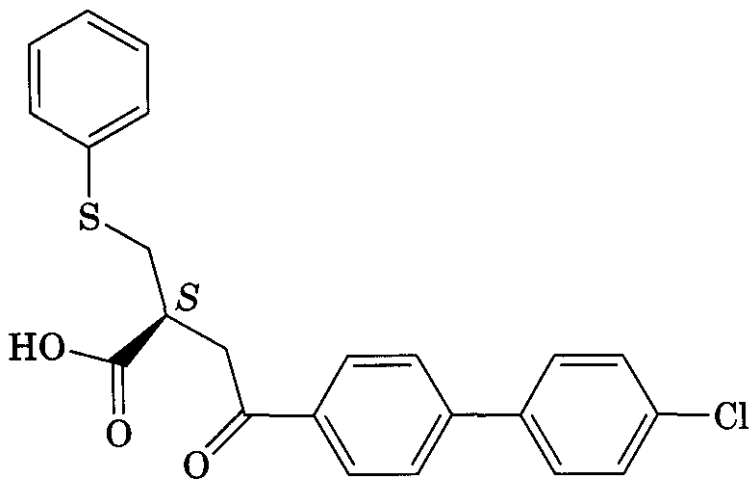
(52) pralnacasan



(53) prinomastat, AG3340



(54) CGS 27023



(55) tanomastat, Bay 12-9566

compounds each have affinities in the nanomolar to picomolar range for several MMPs. The inhibitory profiles and ongoing clinical trials of a variety of drug candidates that inhibit MMPs were reviewed in 2000 (124).

Compound (53) was developed at Agouron through use of SBDD (125) and is under clinical investigation by Pfizer as an anticancer drug and as a treatment of proliferative retinopathy. Compound (54) is a stromelysin inhibitor discovered at Novartis (126), without explicit structural guidance. However, the lead molecule from which (54) was developed was originally obtained by X-ray structure-based inhibitor design targeted against the bacterial zinc-protease thermolysin. Compound (55), with particularly high affinity for the gelatinases, was also developed with consideration of the structures of other MMP–inhibitor complexes, but not through use of iterative SBDD (127). The clinical trials of compounds (54) and (55) have been suspended because of their disappointing efficacy (124). It remains somewhat uncertain which MMP is responsible for specific diseases, and the possibility for biological redundancy suggests that inhibition of several MMPs may be required for treatment of some diseases. SBDD clearly could have a major impact on the discovery of selective MMP inhibitors. These could be useful tools in dissecting the disease relevancy of these targets, as well as providing the selectivity and bioavailability required of effective drugs.

2.5 Oxidoreductases

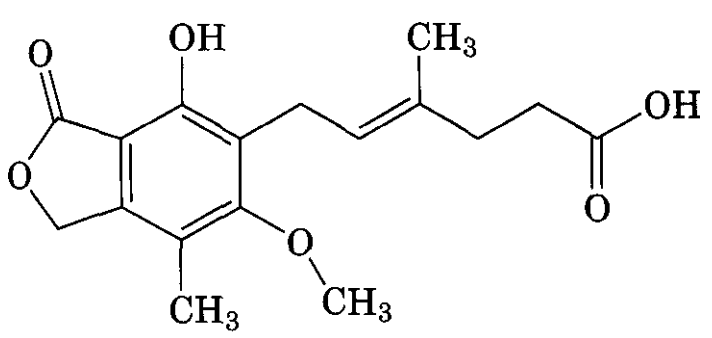
Oxidoreductases catalyze the oxidation or reduction of carbon–carbon, carbon–oxygen, or carbon–nitrogen bonds. Frequently, nicotinamide cofactors are involved, with the oxidized and reduced forms (respectively, NAD^+ or $NADP^+$ and NADH or NADPH) receiving or donating the equivalent of a hydride during this process. Nicotinamide-linked oxidoreductases that have been targeted for the discovery of new therapeutic agents include aromatase, dihydrofolate reductase (mentioned above), aldose reductase, and inosine monophosphate dehydrogenase. SBDD methods have been successfully applied recently to the latter two enzymes to discover agents that are currently

in human testing. The efforts with these two targets are described briefly below.

2.5.1 Inosine Monophosphate Dehydrogenase. Proliferative cells such as lymphocytes have high demands for the rapid supply of nucleotides to support DNA and RNA synthesis, as do viruses during their proliferative phase. The first dedicated step in the de *novo* biosynthesis of guanine nucleotides is conversion of inosinate to XMP, catalyzed by inosine monophosphate dehydrogenase (IMPDH).

$$\text{IMP} + \text{NAD}^+ \rightarrow \text{XMP} + \text{NADH}$$

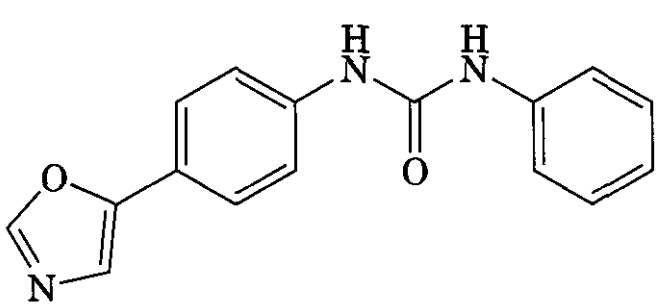
A prodrug form of (56) (mycophenolicacid), a noncompetitive inhibitor of IMPDH, is approved for human therapeutic use as an immunosuppressant (mycophenolate mofetil, CellCept). The use of this drug is hampered by gastrointestinal side effects probably related to the metabolism of the drug. A second class of IMPDH inhibitors is represented by the nucleoside analog mizoribine (also known as bredinin), a prodrug approved for human use in Japan. Such compounds competitively inhibit IMPDH in *vivo* after phosphorylation (128). These drugs validate the strategy of targeting IMPDH for the discovery of immunosuppressants. Other utilities that have been suggested for IMPDH inhibitors are antiviral and anticancer therapies.



(56) mycophenolic acid

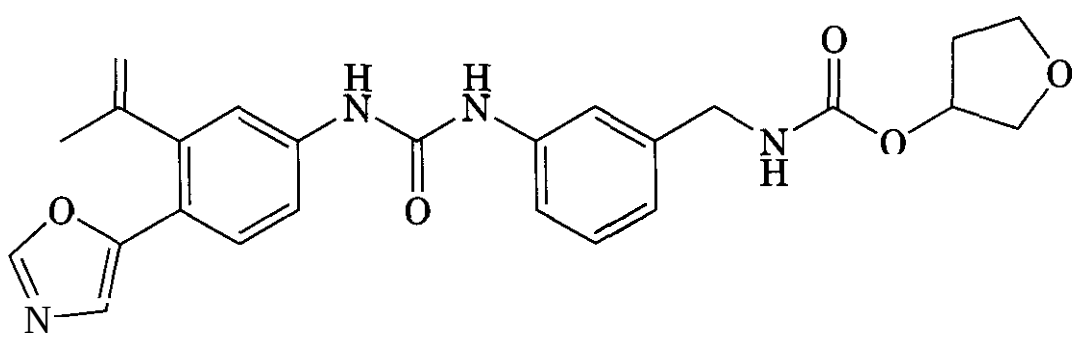
The structure of hamster IMPDH in complex with IMP and (56)was solved at Vertex in the mid-1990s (129). This allowed the visualization of a covalent intermediate, in which a cysteine thiol from the enzyme adds to C2 of the purine ring of the nucleotide substrate. An analogous covalent adduct is postulated to be a key catalytic intermediate during normal turnover (130). The structure was a key tool in the discovery of (57) (VX-497, merimepodip), a novel potent inhibitor of human IMPDH suitable for oral administration (131).

An experimental screen of a diverse library of commercially available compounds for inhibitors of IMPDH identified molecules with the phenyl, phenyloxazole urea scaffold (58) as weak inhibitors. Through use of the computational program DOCK (132), the initial inhibitors were built as models into the experimental structure of the crystalline complex of IMPDH, IMP, and (56). Structural analogs were generated to improve potency in an iterative process, guided by the structural modeling and the observed changes in potency for inhibition of human IMPDH.

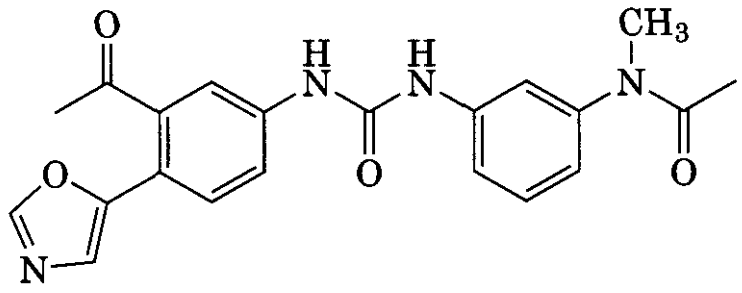


(58)

After this process yielded compound (59), with nanomolar potency, an X-ray structure



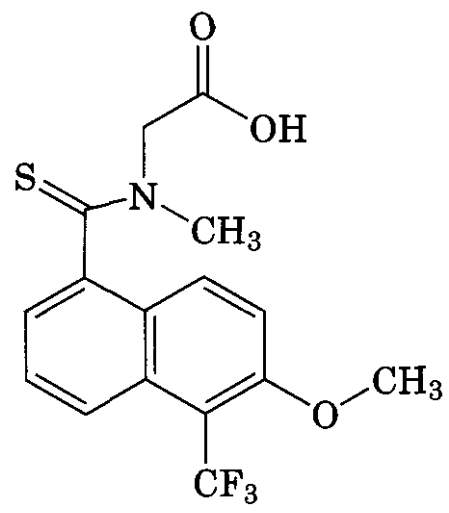
(57) merimepodip



(59)

was determined of (59) bound to the hamster enzyme with IMP. This revealed both similarities and differences between the binding modes of (**56**) and (59). **Aryl** groups of both compounds pack against the covalently tethered purine of the nucleotide. Several hydrogen bonding and hydrophobic interactions with the enzyme are also common between the two inhibitors. However, there are several hydrophobic and van der **Waals** interactions seen in the complex with (59) that are not present with (56). Importantly, the urea moiety of (59) forms a network of hydrogen bonds with an aspartyl carboxylate that is not present in the complex with (56). Further modification of the structure was guided by the X-ray study by use of (**59**), to gain potency in a cell-based assay for inhibition of lymphocyte proliferation. This provided compound (**57**), which Vertex has advanced into clinical trials for treatment of hepatitis C infections.

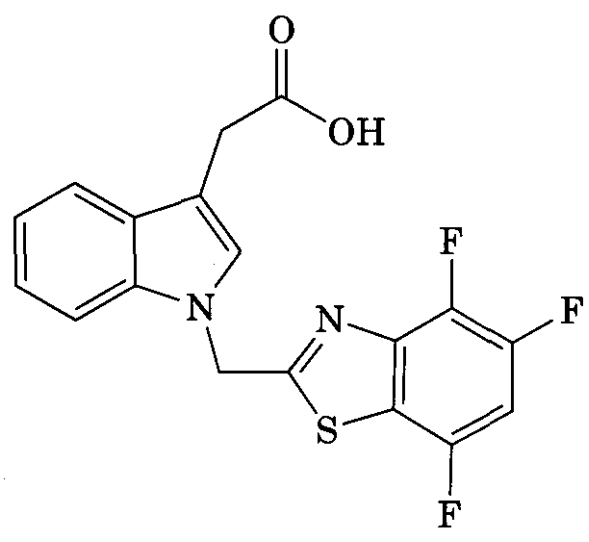
2.5.2 Aldose Reductase. Aldose reductase has been implicated in many of the pathologies resulting from elevated tissue levels of glucose in diabetes mellitus (133, 134). This **nicotinamide**-dependent enzyme catalyzes the conversion of glucose to sorbitol, accumulation of which ultimately results in damage to the eyes, the nervous system, and the kidneys. Given the enormous damage caused by this disease and the difficulty in regulating blood glucose, selective and potent inhibitors of human aldose reductase offer great potential benefit. However, existing drugs that target aldose reductase have unreliable efficacy (135). For example, compound (**60**) (tolrestat) was withdrawn by Wyeth in 1996 because of poor clinical response. Hence, there is still a need to provide an inhibitor of this enzyme that fulfills the potential in the clinic. To minimize the risk of undesired toxicities, clinical



(60) tolrestat

agents that target aldose reductase should not inhibit the closely related aldehyde reductase, an essential hepatic enzyme.

The structure of (60) and other inhibitors bound to porcine aldose reductase (136) provided a rich lode of information on the requirements for potent and selective inhibition of aldose reductase. This was mined by scientists at the Institute for Diabetes Discovery, in a project that began in 1996. The Institute for Diabetes Discovery filed an IND application for (61) (lidorestat, IDD 676), a potent aldose



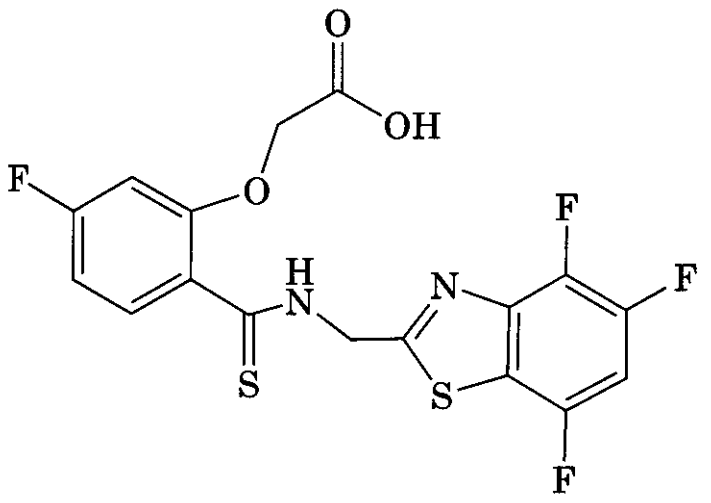
(61) lidorestat

reductase inhibitor, for treatment of diabetic complications, within 30 months of initiating the discovery project on this target. The speed with which this was achieved appears in large part because of the use of SBDD methods.

The X-ray structures showed the cofactor $NADP^+$ buried within the enzyme, with its **C4 redox** center exposed at the bottom of a deep hydrophobic cleft. An anionic binding site is located near $NADP^+$. Several potent **inhibi-**

tors bind within the hydrophobic cleft and interact with the anionic site. The binding of potent inhibitors induces a conformational change, opening an adjacent hydrophobic pocket. The conformation induced by (**60**) differs from that caused by other, less selective inhibitors. This "specificity" pocket was thought to offer an opportunity for selective inhibition of aldose reductase while sparing aldehyde reductase. Hence, this structural study provided an initial pharmacophore for both potency and selectivity.

The SAR for this pharmacophore was developed with a series of synthetically accessible salicylic acid derivatives that were scored for potency and selectivity with the purified enzymes, and efficacy in a diabetic rat model **(137).** One of the most potent and selective of the derivatives was (**62**), containing the benz-



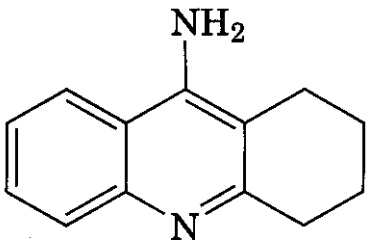
(62)

thiazole heterocycle. The SAR was employed, guided by the structures of selected inhibitor complexes, to design a novel indole scaffold to present the pharmacophoric elements (M. **Van** Zandt, personal communication). The optimization of this series provided the clinical candidate (**61**) (138).

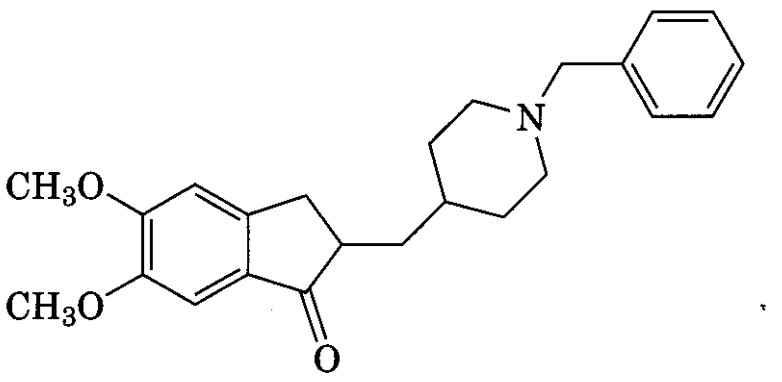
2.6 Hydrolases

Some other hydrolytic enzymes, in addition to proteases, that are important drug targets include protein phophatases, phosphodiesterases, nucleoside hydrolases, acetylhydolases, glycosylases, and phospholipases. Structure-based inhibitor design is currently being applied to a number of these enzymes. The last three mentioned have been successfully targeted in SBDD projects that have produced compounds that are either launched or in clinical trials.

2.6.1 Acetylcholinesterase. A pronounced decrease in the level of the neurotransmitter acetylcholine is one of the most pronounced changes in brain chemistry observed in the sufferers of Alzheimer's disease (139). Several drugs that are approved for the treatment of the dementia thought to result from this neurotransmitter deficit act by inhibiting acetylcholinesterase. These include (63) (tacrine, or

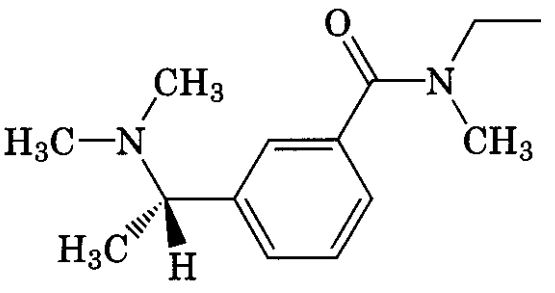


(63) tacrine



(64) donepezil

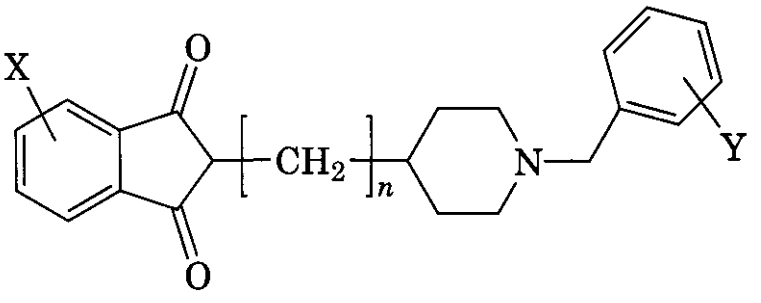
Cognex, a Pfizer drug that was the first such agent approved for this indication), (64) (donezepil), and (65)(rivastigmine). Several other agents are in clinical trials. Disappointing efficacy is observed with the existing drugs, arising from dose limitations that are likely attributable to the inhibition of acetylcholinesterase



(65) rivastigmine

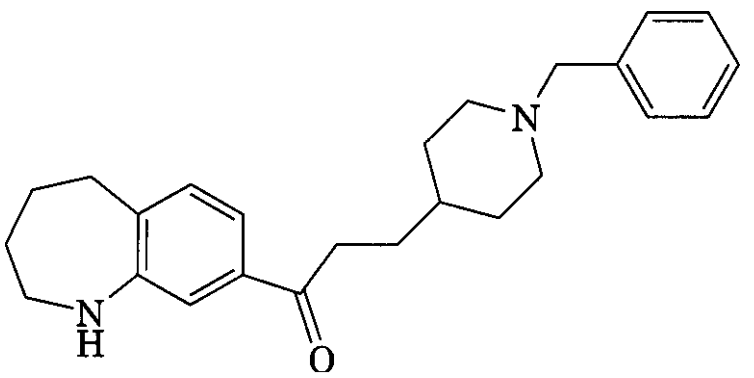
in peripheral tissues (140). This may be a consequence of the high serum levels required to get these highly cationic molecules to penetrate the blood–brain barrier.

In a discovery project that is reminiscent of the discovery of captopril, scientists at Takeda created a hypothetical structure for the active site of acetylcholinesterase, based on *SAR* from previous biochemical and medicinal chemical work (141). The model consisted of (in addition to the serine protease-like catalytic machinery) an anionic binding site separating two discrete hydrophobic binding sites. This model was then used to design inhibitors of the enzyme (reviewed in ref. 142). One set of analogs examined were based on the *N*-(ω-phthalimidylalkyl)-*N*-(ω-phenylalkyl)-amine (scaffold 66). An iterative process of testing,



(66)

analysis, design, and synthesis, by use of this and closely related scaffolds, resulted in the production of (67) (TAK-147), which is currently in clinical trials for treatment of the dementia resulting from Alzheimer's disease (142).

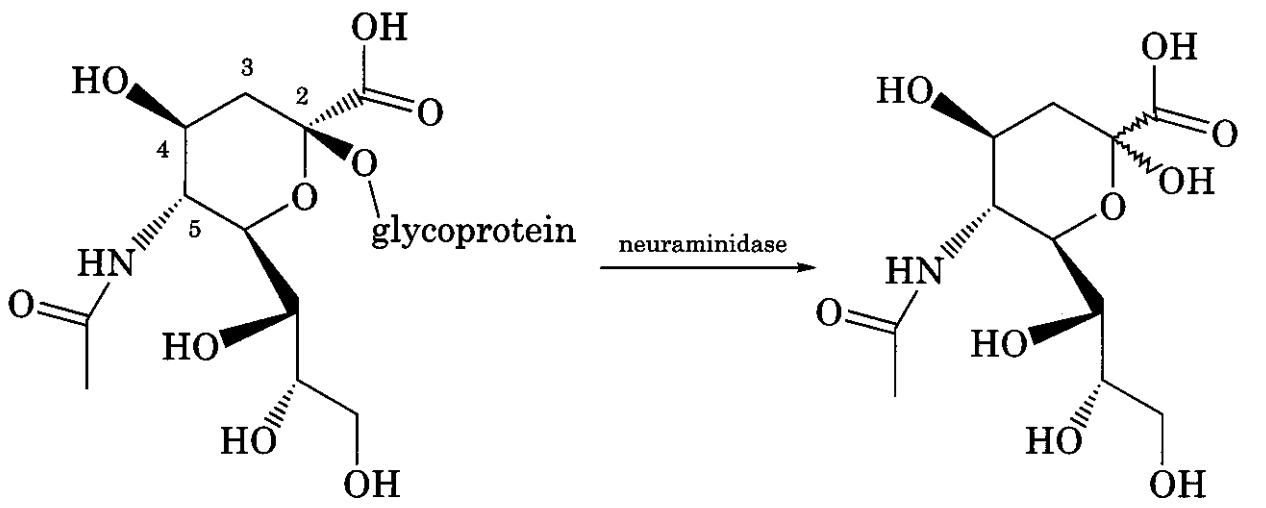


(67) TAK-147

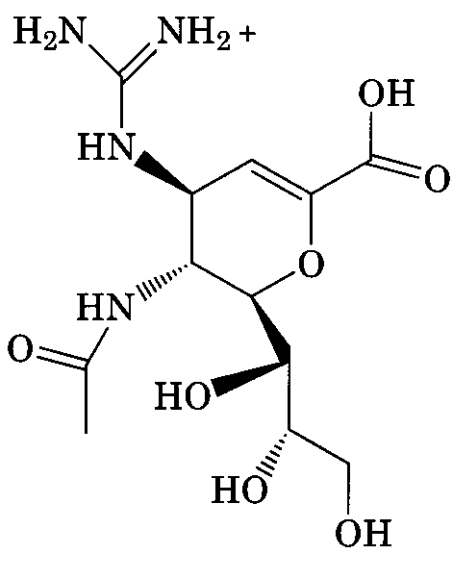
The design of (66) was partially based on the structures of previously known inhibitors. The two aryl substituents were intended to bind to the hydrophobic binding sites, placing the central amine cation into the anionic binding site. The length of both alkyl linkers was varied, and the effect of adding a third alkyl substituent was examined. The phthalimide portion of the structure was chosen to improve the synthetic accessibility of the analogs needed for this exercise. The compounds were tested not only for inhibitory potency toward rat cerebral acetylcholinesterase, but also for peripheral response and toxicity in dosed intact rats. After the work was under way, Sussman and coworkers solved the atomic structure of acetylcholinesterase from the electric eel, including complexes with several inhibitors, by X-ray crystallography (143). The availability of this structure made it possible to retrospectively analyze the basis for the *SAR* in this series of compounds, by use of DOCK (144).

2.6.2 Neuraminidase. Influenza virus infections cause severe human suffering throughout the world and economic damage in the billions of dollars annually, although some years are worse than others. In 1918 a pandemic caused by this disease killed an estimated 40 million people (145). An important protein in the infectious process is the viral neuraminidase, an integral membrane protein whose catalytic domain is exposed on the viral surface. Neuraminidase catalyzes the hydrolytic cleavage of sialic acid (68, *N*-acetylneuraminic acid) from glycoproteins and extracellular mucin on the surface of the host cell. A different viral surface protein tightly binds to terminal sialic acid residues. which promotes the initial infection, but prevents release of viral progeny from the host cells, unless and until the terminal sialic acids are hydrolytically cleaved by viral neuraminidase. Thus, neuraminidase enables the infection to propagate.

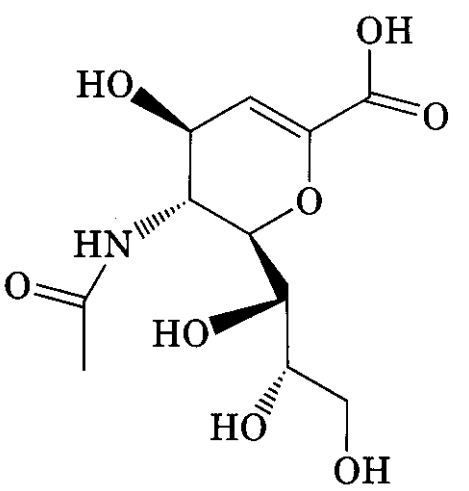
The first X-ray structure of influenza neuraminidase was determined in the early 1980s (146). Ten years later, a landmark paper (147) described a highly efficient drug design project at Monash University in Australia. This project yielded antiviral compound (69)(zanamivir, Relenza, or Flunet), which was developed into one of the first drugs to be created through use of SBDD. Previous structural work had revealed that the active site of neuraminidase has several rigid pockets and nu-



(68) sialic acid



(69) zanamivir



(70) Neu5Ac2en, DANA

merous charged groups. Electrostatic interactions significantly affect the conformation of bound sialic acid, which is deformed into a high energy conformer, attributed in part to the interactions between the 1-carboxylate and arginine side-chains of the protein. This deformation may play a key role in catalysis. Synthesis of a sialic acid analog that is dehydrated across the C2–C6 bond of (**68**) had provided the putative transition state mimic (**70**) (sometimes referred to as Neu5Ac2en, or as 2-deoxy-2,3-dehydro-*N*-acetylneuraminic acid, DANA).

Compound (70) inhibits neuraminidase with micromolar potency (148). Examination of the binding mode of (70) in the active site of neuraminidase (Fig. 10.16) led to the replacement of the 4-hydroxyl by cationic groups, first an amino and then a guanidino group (147). These groups strongly interact with anionic amino acid side chains (corresponding to Glu120 and Glu229 shown in Fig. 10.16) in the neuraminidase active site. In the case of the guanidine substitution, the binding affinity for neuraminidase was increased about 5000-fold and provided (**69**), which inhibits viral release in cell cultures and decreases the severity of influenza virus infections in humans. Subsequently, the X-ray structures of neuraminidases from several different influenza subtypes complexed with (69) were analyzed (149). Although the positions of protein residues were well conserved, the water structure seen in these different complexes was quite variable. This may explain the varying potency of (69) against different strains of virus.

One problem with (69) is that it is not well absorbed by an oral route, and so must be administered as an aerosolized powder inhaled into the virus-infected lungs. Two other neuraminidase inhibitors with nanomolar affinities (71 and 72) have been developed through the use of SBDD methods to yield orally bioavailable drugs. The development of these agents was facilitated by the fortuitous discov-

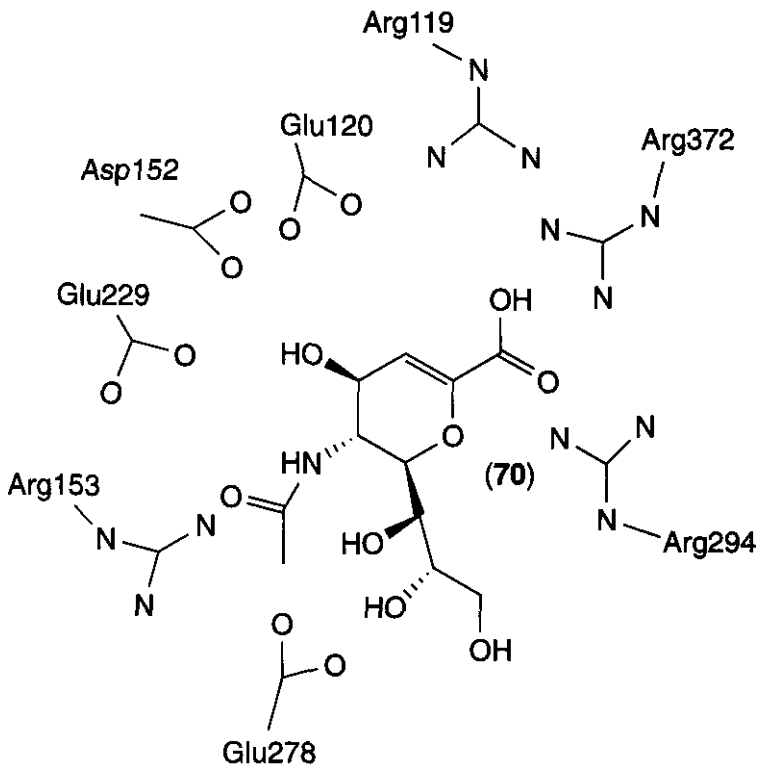
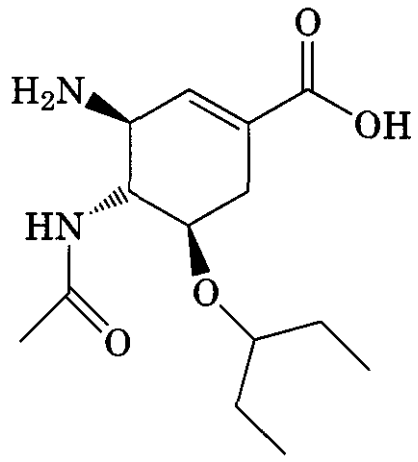
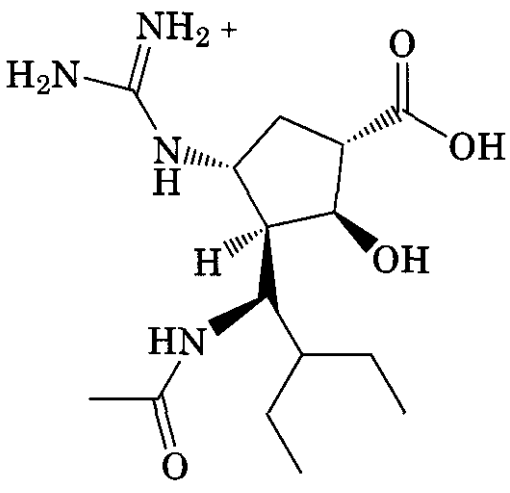


Figure 10.16. View from above: Polar **amino** acid side-chains surrounding (**70**), bound into the **active** site of influenza virus neuraminidase (Scheme 10.1 based on PDB code **1NNB**, the coordinates of an X-ray structure described in Ref. 148).



(71) GS 4071



(72) BCX 1812

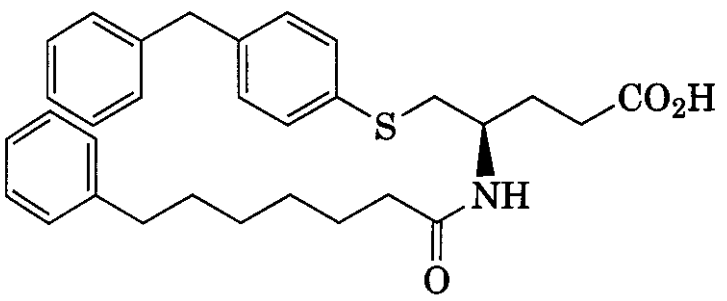
ery by scientists at Biocryst, that analogs of (**69**) in which the cyclic scaffold is a phenyl moiety are much more potent inhibitors if they lack the glycerol side chain! This was subsequently discovered by X-ray structural studies to be attributed to the creation of an unanticipated hydrophobic pocket upon rearrangement of the **Glu278** side chain carboxylic acid, which forms several hydrogen bonds with the glycerol portion of (69) (Fig. 10.16).

Replacement of the permanently **cationic** guanidine by an **amine** (71) promoted better intestinal absorption, but also greatly decreased the **affinity** for neuraminidase. Structure-guided modification of the carbocycle's substituents was used to recover this lost potency. Compound (71) (GS 4071) was developed by Gilead Sciences (150). The ethyl ester of (71) is a **prodrug** (oseltamivir or GS 4104) that has been approved for oral dosing to treat influenza infection. Another amphiphilic **carbocycle**, compound (72) (peramivir, RWJ-270201, or BCX 1812) was developed by **BioCryst** (151) through use of SBDD, and is in clinical trials. The use of clever synthetic routes, biochemical assays for neuraminidase inhibition, a mouse infection model, and X-ray structural information were all valuable tools in the development of both (71) and (72). Optimization of the **affinity** required the examination of a variety of alkyl substituents in bdth cases, to exploit the new hydrophobic pocket created by the conformational change primarily involving **Glu278**. The ability of the **cyclopentyl** ring in (72) to replace the **six-membered** ring illustrates that differing central scaffolds can display the essential interacting groups in an effective way.

2.6.3 Phospholipase A2 (Nonpancreatic, Secretory). Phospholipases A2 (**PLA2s**) are a diverse family of hydrolases that cleave the ***sn-2*** ester bond of phospholipids. The fatty acid produced is frequently arachidonate, the precursor to the proinflammatory eicosanoids. In several human inflammatory pathologies (**e.g.,** septic shock, rheumatoid arthritis), a nonpancreatic secretory form of **PLA2** (**hnps-PLA2**) is present in extracellular fluids at levels many-fold higher than normal (**152**). The design of bioavailable inhibitors of this Ca^{2+}-dependent **isoform** of **PLA2** as inflammatory

drugs is therefore an attractive goal (153). To be an effective drug, such an inhibitor would also need to be selective for hnps-PLA2 vs. the closely related pancreatic PLA2. Whether selectivity is needed against the quite different cytosolic PLA2 is unclear.

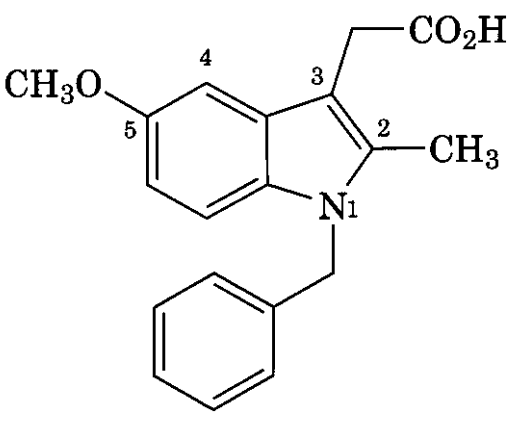
Investigators at an AstraZeneca laboratory (previously Fisons) have used multidimensional NMR and computational techniques to develop an active site model for cytosolic PLA2 (154, 155). Synthesis of compounds based on this model led to (73) (FPL-67047), reported



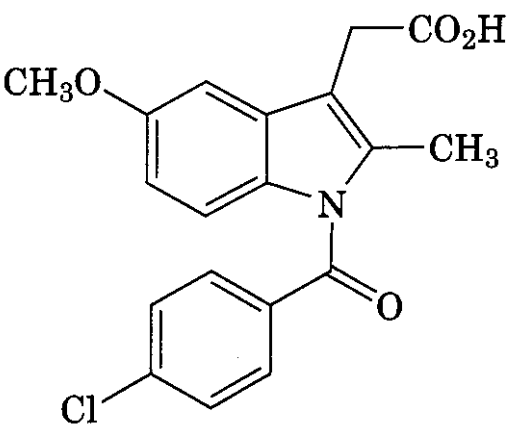
(73) FPL-67047

to be a development candidate for treatment of inflammation (156).

Investigators at Eli Lilly began a project to develop PLA2 inhibitors by investing the effort to clone, overproduce, purify, crystallize, and determine the structure of hnps-PLA2 (157). This also provided the reagent needed for a massive screening campaign to identify hnps-PLA2 inhibitors. They were thus prepared to apply SBDD methods when the screening of Lilly's small molecule collection yielded a weak inhibitor. The hit (74) was surprisingly similar to indomethacin (75), a nonsteroidal anti-inflammatory drug that acts by inhibiting cyclooxygenase.

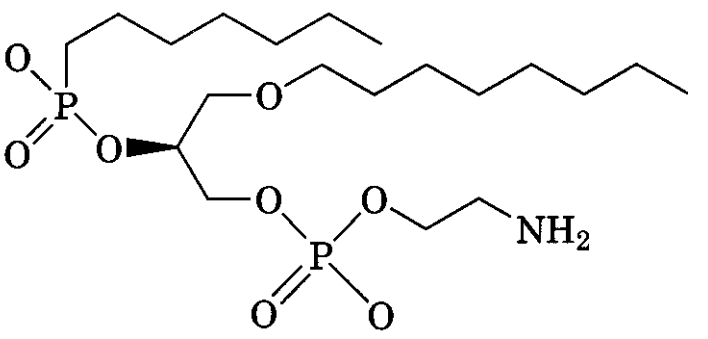


(74)



(75) indomethacin

The crystal structures of recombinant hnps-PLA2 bound to (74) and (75) were solved (158), and compared with the previously known structures of PLA2s complexed with substrate mimics (159, 160), including the phosphonate-containing transition state analog (76). The earlier structures revealed several key features. These were: (1) the filling of a significant hydrophobic crevice, (2) the displacement (by the sn-2 alkyl moiety) of the His6 side-chain into a solvent-exposed position to create an adjacent cavity, (3) the coordination of the active site calcium, and (4) formation of hydrogen bonds to His48 and Lys69. The polar contacts were provided by the nonbridging phosphate and phosphonate oxygens in the complex with (76).

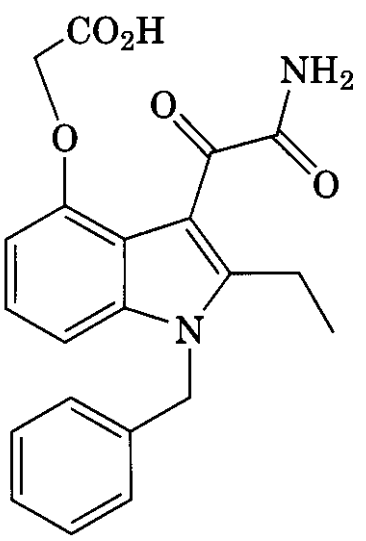


(76) hnps-PLA2 transition-state analog

The screening hit (74) bound in the hydrophobic crevice, similarly to the substrate mimics, with the 1-benzyl moiety of (74) bound in the adjacent cavity and displacing the enzyme's His6 imidazole. However, there were two surprising findings. First, despite the presence of 10 mM calcium in the crystallization liquor, there was no bound calcium, an essential active-site component, although weakly binding (K_d = 1.5 mM). Second, the carboxylic acid of (74) formed a hydrogen

bond with another active-site acid, the side chain of Asp49. The latter finding again emphasizes the importance of experimental structures to guide improvements of inhibitor potency, given that placing two presumed anions so close together would likely never have been predicted by a computational model. Other slight conformational changes were observed to accommodate the 5-methoxy group of (74).

The inhibitor's 3-acetate moiety was converted to an acetamide in a successful attempt to restore the active site calcium, form a hydrogen bond to His48, and increase potency. The crystal structure of the complex with the amide version of (74) also revealed a significant reorientation of the indole core and 5-methoxy substituent, resulting in an unanticipated 5-Å movement of the terminal methyl. Further changes in inhibitor structure were guided by iterative structural studies and functional assays of potency and selectivity. These changes involved the use of substituents at positions 3 or 4 to optimize coordination of the metal ion, extension of the van der Waals interaction by lengthening the 2-methyl to an ethyl, and conversion of the 3-acetamide to glyoxamide (159,161). This resulted in the synthesis of (77) (compound LY315920), which has 6500-fold greater affinity for hnps-PLA2 than did the original hit molecule (74). LY315920 effectively inhibits hnps-PLA2 in the serum of transgenic mice dosed with the compound orally or i.v., and is undergoing clinical trials in the United States and Japan (162,163).

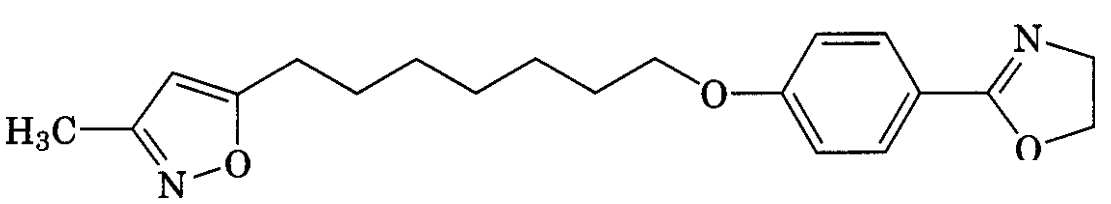


(77) LY315920

2.7 Picornavirus Uncoating

Picornaviruses, which include the rhinoviruses and enteroviruses. are RNA viruses that cause several infectious human diseases. These diseases include common colds as well as life-threatening infections of the respiratory and central nervous systems. Effective treatments of these diseases would relieve much human suffering, save many lives, and have great economic benefit. There are over 100 serotypes of rhinoviruses alone, making it impossible to generate a vaccine effective against infections by all variants of the virus (164).

The Achilles heel of picornaviruses has been suggested to be that part of the virus structure that interacts with the cell surface receptor because those structural features must be well conserved (165). The virus particle consists of a positive-strand RNA coated by an icosahedral shell, containing 60 copies of four distinct β-barrel proteins (166). These structural proteins contain the binding site for the cellular receptor and undergo significant conformational changes to liberate the viral RNA genome during infection of the cell. A series of isoxazoles that inhibit this picornavirus "uncoating" process were discovered in the early 1980s by scientists at Sterling Winthrop, by use of an in vitro cellular assay for antiviral activity (167–170). One of these, compound (78) (WIN-51711, disoxaril), gave a 50% suppression of viral plaque formation in this assay at 0.3 μM. Compound (78) was also effective in animal models (171) and entered phase I clinical trials, but failed to advance because of its toxicity. Compound (78) was shown (172) to bind to viral capsid protein



(78) WIN-51711, disoxaril

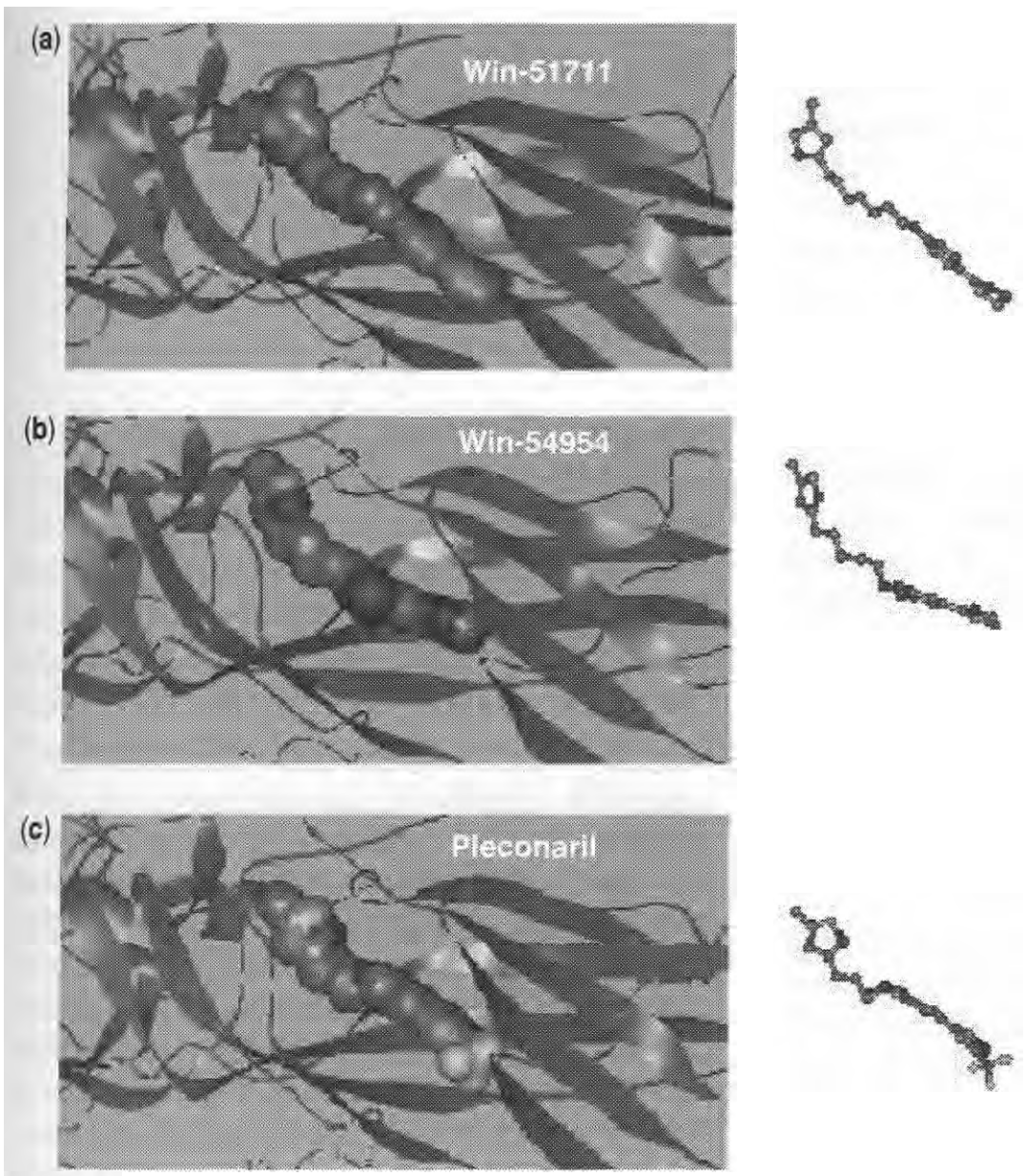
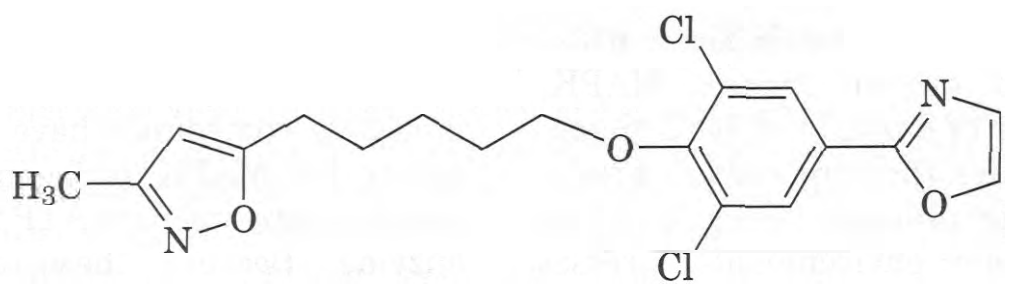


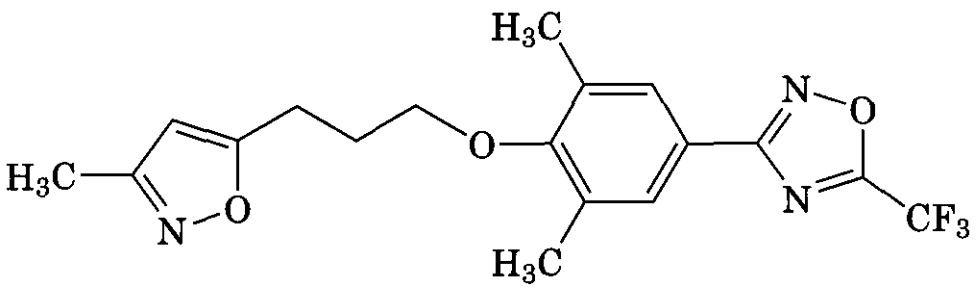
Figure 10.17. Structure of rhinovirus capsid protein VP1 showing the bound conformation of antiviral isoxazole compounds (**78**) [disoxaril, WIN-51711: panel a, top], (79) [WIN-54954: panel b, middle], and (**80**) [pleconaril, WIN-63843: panel c, bottom]. The PDB codes for the X-ray structural model coordinates used to create these views are: 1PIV (for **78**), 2HWE (for **79**), and 1C8M (for 80). On the left side of each panel, the inhibitors are shown as van der Waals surfaces, and the protein as a ribbon diagram. On the right side, the structures of the inhibitor alone are shown, from the same view, as ball and stick representations. See color insert.

VP1, within a hydrophobic pocket in the floor of the "canyon" that contains the binding site for the cell surface receptor (Fig. 10.17A). Structural changes induced in the canyon floor upon binding of such molecules may also inhibit receptor binding directly (173). X-ray crystallographic studies of (**78**) and analogs bound to the target protein VP1 were an essential part of the iterative optimization process that led to safer and more effective antiviral agents (174–176). The goal of the process was to generate a compound that is potent, chemically and metabolically stable, and effective against as many serotypes of the virus as possible. There was therefore a need to balance potency and selectivity, and the structural information helped to guide compound design in pursuit of this balance.

A second-generation compound, (79) (WIN-54954) also advanced into clinical tests, but had disappointing efficacy in Phase II trials, probably because of extensive metabolism. Modification of the phenylisoxazole, guided by both structural and metabolic considerations (177), allowed the creation of a stable and potent antiviral, the third-generation compound (80) (WIN-63843, pleconaril, or Picovir) (178). This compound was evaluated in Phase III clinical trials and showed efficacy in humans. Oral dosing of virally infected patients with



(79) WIN-54954



(80) VP63843, **pleconaril**

(80) three times daily decreased the average time needed to become free of cold symptoms from 10 days to between 8 and 9 days, and also reduced the duration of severe cold symptoms from 4.5 to 3.5 days (179). During the clinical studies to support the new drug application for (**80**), about a quarter of the clinical isolates (of rhinovirus present initially or during the treatment) were resistant to the compound. The majority of these resistant viruses had a single mutation at VP1 residue Ile98, which directly interacts with (80) bound to VP1 in wild-type virus. The clinical data also showed the elevation in some patients of hepatic cytochrome P450 levels during treatment with (**80**), raising concerns about potentially hazardous drug–drug interactions. ViroPharma sought and failed in early 2002 to gain the approval of the U.S. Food and Drug Administration for its new drug application for (80) for treatment of the common cold.

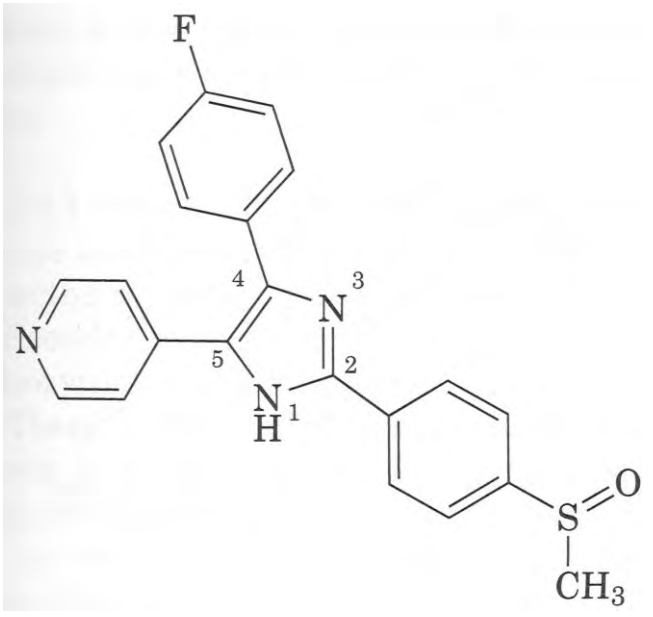
2.8 Phosphoryl Transferases

Protein kinases and phosphatases play vital roles in intracellular signaling pathways and in the integration and control of major cellular processes. Kinases and other phosphoryl group transferases are essential in the metabolism of lipids, nucleotides, and other small biomolecules. The use of SBDD methods on such targets has expanded as more of their X-ray structures have been solved, and will continue to grow as more targets are validated for their involvement in human diseases.

2.8.1 Mitogen-Activated Protein Kinase p38α.

Mitogen-activated protein kinase (MAPK) p38α is a member of a family of Ser/Thr-specific protein kinases that are activated upon exposure of cells to mitogens such as bacterial lipopolysaccharide or environmental stresses such as exposure to UV irradiation or chemical oxidants. MAPK p38α has a central role in integrating the inputs from a complex signaling network. Activation of MAPK p38α requires the dual phosphorylation of conserved threonine and tyrosine residues on a loop near the enzyme's active site (180). The unactivated (nonphosphorylated) enzyme has a very low affinity for ATP, but can bind to pyridinylimidazole inhibitors (181, 182). The activated enzyme in turn phosphorylates numerous substrates, including several transcription factors. This leads to activation of the transcription of many genes and causes the release of proinflammatory cytokines, primarily interleukin-1β (IL-1β) and tumor necrosis factor (TNFα). MAPK p38α was identified as a central player in this inflammatory pathway in a key study by scientists at SmithKline Beecham (183). The study involved the molecular cloning of the genes encoding proteins that bind to anti-inflammatory pyridinyl-imidazole compounds already known to block the biosynthesis of IL-1β and TNFα. The binding proteins turned out to be members of a known kinase family. Since this finding, the enzymes in the MAPK pathway, and especially MAPK p38α, have been attacked by many scientists seeking to discover anti-inflammatory drugs (184).

Compound (81) (SB 203580), a specific inhibitor of MAPK p38α, is a prototype for the pyridinyl-imidazole compounds (185). This compound is active in animal models of several inflammatory diseases (186), but was not itself pursued as a clinical candidate because of its inhibition of other enzymes, including hepatic cytochrome P450 reductases. The pyridinyl-imidazole compounds have dissociation constants for MAPK p38α in the nanomolar range, competing with ATP for binding to the enzyme. Because these compounds bind tightly to the unactivated enzyme, which has a



(81) SB203580

low affinity for ATP, they are able to compete effectively even *in vivo*, where the ATP concentration is in the millimolar range. The X-ray structures of (81)and several other pyridinyl-imidazole compounds in complexes with human MAPK p38α were solved in a collaborative effort between scientists at SmithKline Beecham and the University of Texas (187). Several X-ray structures of human MAPK p38α with and without bound inhibitors have also been solved by scientists at Vertex (181, 188).

The structures of the inhibited enzyme were useful in understanding what parts of the compounds were responsible for strong binding to MAPK p38α. As shown in Fig. 10.18, both hydrophobic and hydrogen bonding interactions are important components of the inhibitor binding pocket. This structure suggested that Thr106 is an important structural determinant of the selective inhibition of MAPK p38α by the pyrimidyl-imidazoles, which have low affinity for other closely related kinases. Mutation of Thr106 results in the loss of sensitivity to these inhibitors, whereas the replacement of the corresponding residue in another kinase (ERK2) by threonine caused the mutated variant to become sensitive to these inhibitors (189, 190).

The X-ray structural models were also used at both SmithKline Beecham (later GlaxoSmithKline) and Vertex to guide the design of new inhibitors. For example, both N1 of the central imidazole and the 2-(para-methylsulfonyl)-phenyl substituent in enzyme-bound (81)face a channel that opens to bulk solvent. This observation led to the design of (82) (VK19911) at Vertex (181) and (**83**) (SB242253) at GlaxoSmithKline (191). Compound (83)is fivefold more potent than (81)*in vivo*, in a mouse disease model, and was advanced into human clinical trials for treatment of rheumatoid arthritis (192). The piperidine on N1 of (82) and (83)was designed to form a salt bridge with Asp168. This interaction, and the preservation of other binding interactions, was directly demonstrated (181) for compound (**82**).

Analysis of the structural information from the X-ray models allowed the design at Vertex of a new scaffold for potent inhibition of MAPK p38α, as shown for compound (84) (VX-745). This design process, SBDD through

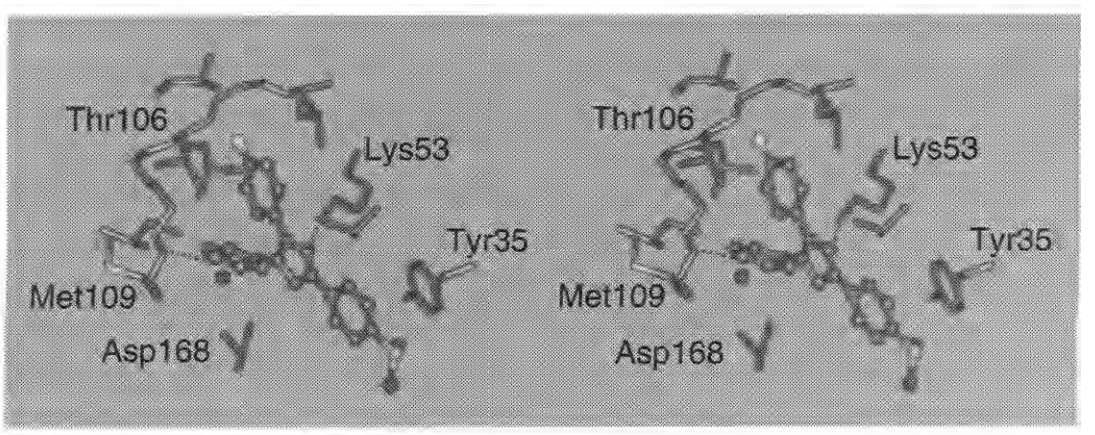
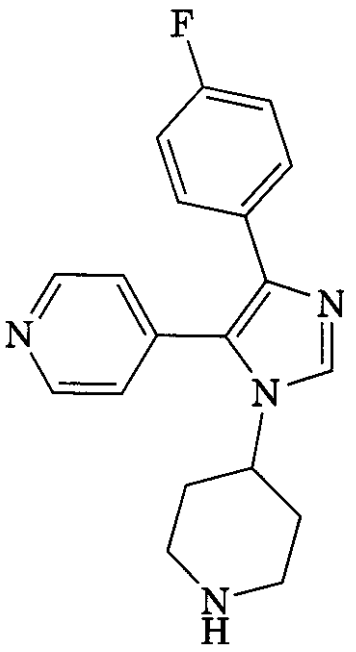
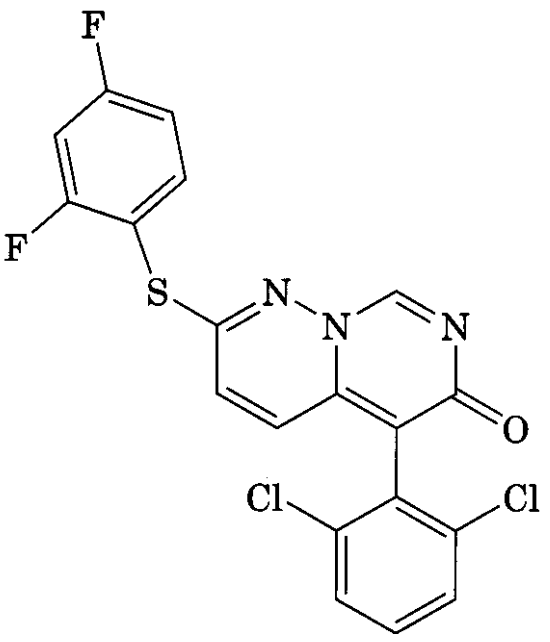


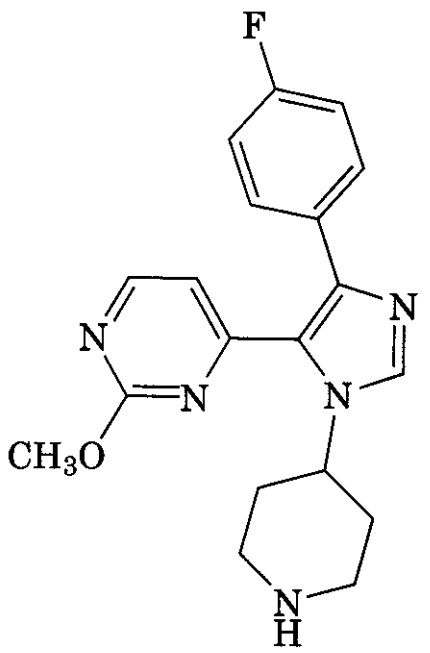
Figure **10.18.** Binding of SB203580 (shown as a ball and stick structure) in the active site of MAPK p38α. In addition to the side chains of the labeled residues, the protein backbone between Leu104 and Met109 is shown, as well as several aliphatic side chains and a water molecule (red sphere). Hydrogen bonds (dotted lines) are shown between the backbone amide of Met109 and the inhibitor's pyrimidinyl nitrogen, and between the ϵ-amino of Lys53 and the inhibitor's imidazole N3. This figure is based on the PDB coordinate set 1A9U (187). See color insert.



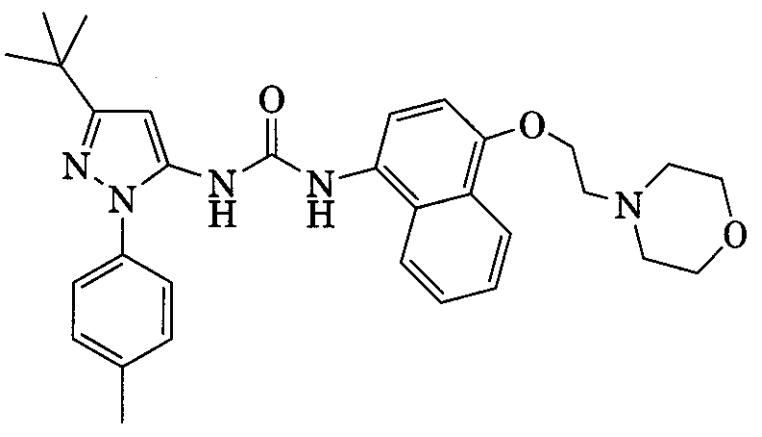
(82) VK19911



(84) VX-745



(83) SB242253



(85) BIRB-796

use of a crystal structure of **MAPK p38α** to design potent inhibitors with potential utility as human therapeutics, is the subject of an international patent application by Vertex, published in 2000 (193). The binding mode for (84) has not been disclosed, but the compound was advanced into clinical trials (194). Vertex has since discontinued the clinical trials of (**84**) because of the potential for toxicity, based on animal data, but in mid-2002 Vertex began a phase I clinical study of a new compound targeted against **MAPK p38α**.

Scientists at Boehringer Ingelheim recently described (195, 196) their discovery of an orally active inhibitor of **MAPK p38α**, compound (85) (**BIRB-796**), that is very different from earlier inhibitors. This compound, whose K_d for **MAPK p38α** is 100 picomolar, has entered phase II clinical trials for treatment of rheumatoid arthritis. The lead compound that led to compound (85) was a **diaryl** urea originally identified by high throughput screening. X-ray structural studies revealed novel modes of binding for both the lead compound and (85) in the active site of **MAPK p38α**. Their binding sites are adjacent to the active site but do not directly overlap with that of ATP; rather, their binding mode changes the conformation of **MAPK p38α** such that ATP cannot bind. The optimization of the lead compound to clinical candidate (85) was an iterative process using clever synthetic chemical design, biochemical assays for affinity, X-ray crystallographic studies of key complexes, and cell-based and animal models. The development of (**85**) as a **MAPK p38α** inhibitor with

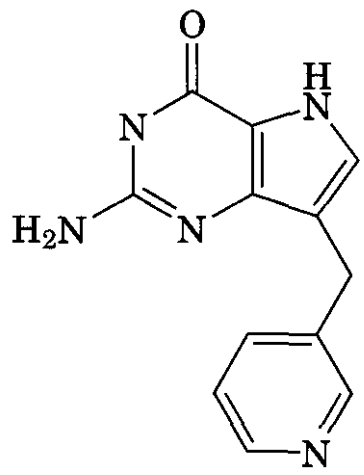
efficacy in vivo makes it evident that there are multiple ways to effectively inhibit this enzyme.

2.8.2 Purine Nucleoside Phosphorylase. Purine nucleoside phosphorylase (PNP) catalyzes the reversible phosphorolysis of purine nucleosides to the purine base and ribosyl or 2-deoxyribosyl-a-1-phosphate.

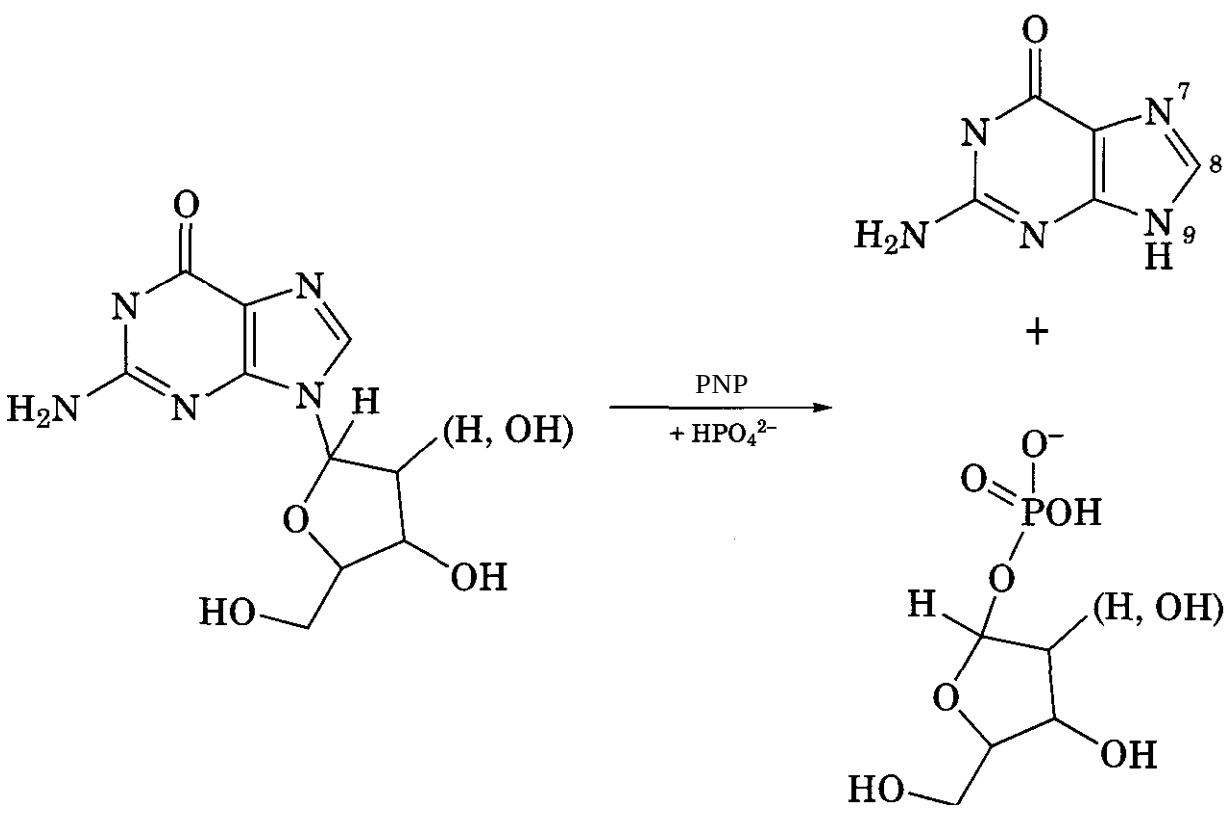
The vital role of PNP in the proliferation of T-cells is evident from the fact that people with an inherited deficit in this activity have 30- to 100-fold lower numbers of T-lymphocytes than normal (197). The accumulation of dGTP and the resulting inhibition of ribonucleotide reductase in PNP-deficient T-cells causes the suppression of T-cell proliferation. B-lymphocytes are unaffected. Hence, small molecule inhibitors of PNP could be used to treat T-cell lymphomas and other T-cell–mediated diseases such as psoriasis. Adjunct therapy with PNP inhibitors could also block the catabolism of therapeutically useful nucleoside analogs.

Human PNP is a homotrimer of 32-kDa subunits. The X-ray structures of the apoenzyme and some substrate analog complexes were described in 1990. Each of the three identical active sites, located near the subunit interfaces, are composed primarily of residues from one subunit, with Phe159 participating in the active site of the adjacent subunit (198).

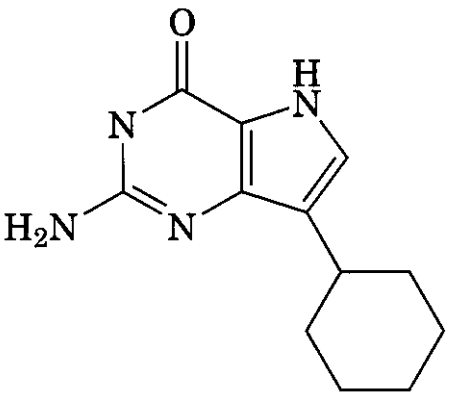
Scientists at BioCryst, CIBA–Geigy, Southern Research Institute, and the University of Alabama collaborated to design inhibitors of human PNP (199,200). The project used an iterative process, in which new compound design was guided by synthetic considerations, computer graphics analysis of X-ray structural models, computational (Monte Carlo and energy minimization) methods, and the inhibitory potency of the compounds against PNP in vitro. Evaluation of the most potent inhibitors by use of cell-based assays, followed by pharmacokinetic and pharmacological characterization of several inhibitors in animal models, led to the choice of (**86**) for advancement into clinical trials. Compound (**86**) (BCX-34, peldesine) is being evaluated for treatment of psoriasis and skin cancer (201,202).



(86) peldesine



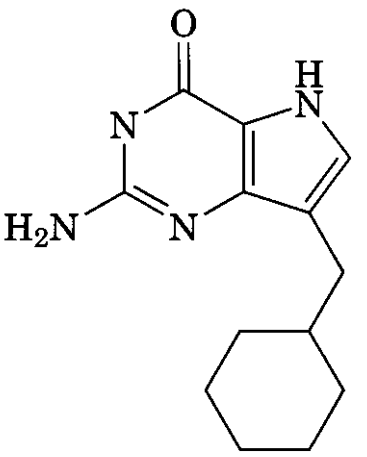
In the SBDD project that produced (86), the design work was initiated through use of the X-ray structure of the PNP apoenzyme, but was more successful when the structures of the PNP–guanine complex and other complexes were available (199). The PNP–guanine crystal structure showed no important interactions with N9, and indicated a potential for hydrophobic interaction in the vicinity of the substrate ribose (Fig. 10.19). To test this, the 9-deaza compound (87) was synthe-



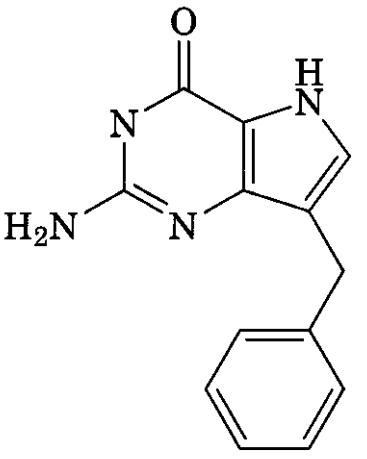
(87)

sized. This was a weak PNP inhibitor (measured $IC_{50} = 1\ \mu M$).

The X-ray structure of the complex between PNP and (87) showed that the hydrophobic interaction dominated the binding mode, and resulted in the disruption of the hydrogen bonding interactions seen in the guanine complex (i.e., Fig. 10.19). To increase the spacing between the hydrophobe and the purine mimetic, compounds (88) and (89) were made. These had affinities for PNP in the nanomolar range. X-ray crystallographicanalysis indicated new hydrogen bonding interactions with these 9-deaza compounds (shown for 89 in Fig. 10.20), made possible because N7 is protonated.



(88)



(89)

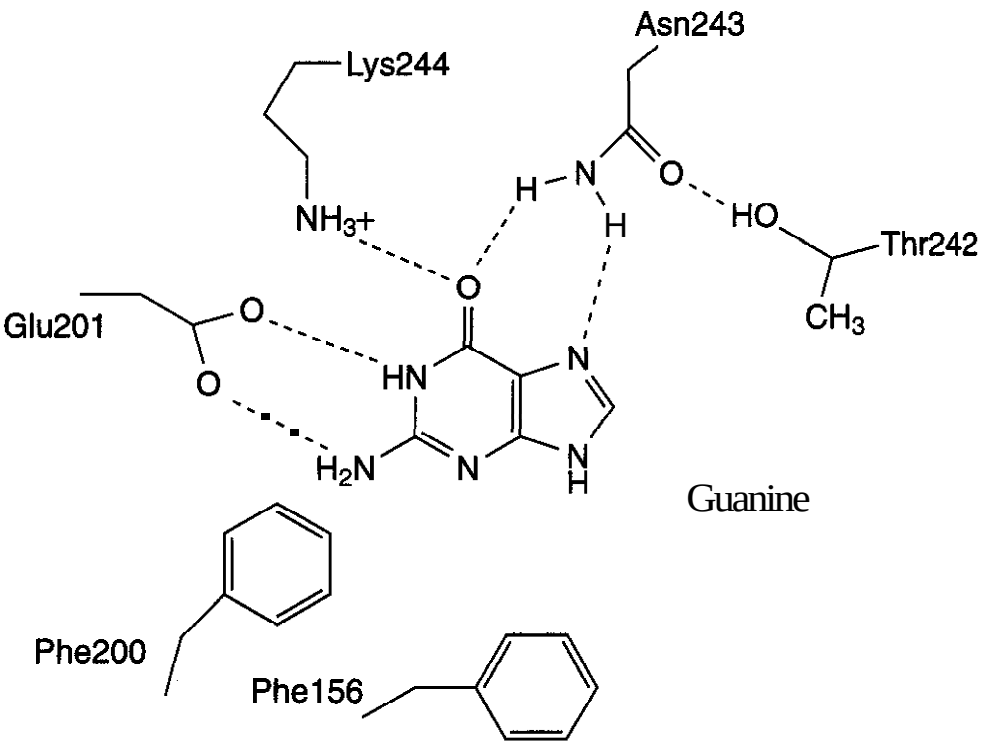


Figure 10.19. Binding interactions in the active site for the complex between guanine and PNP.

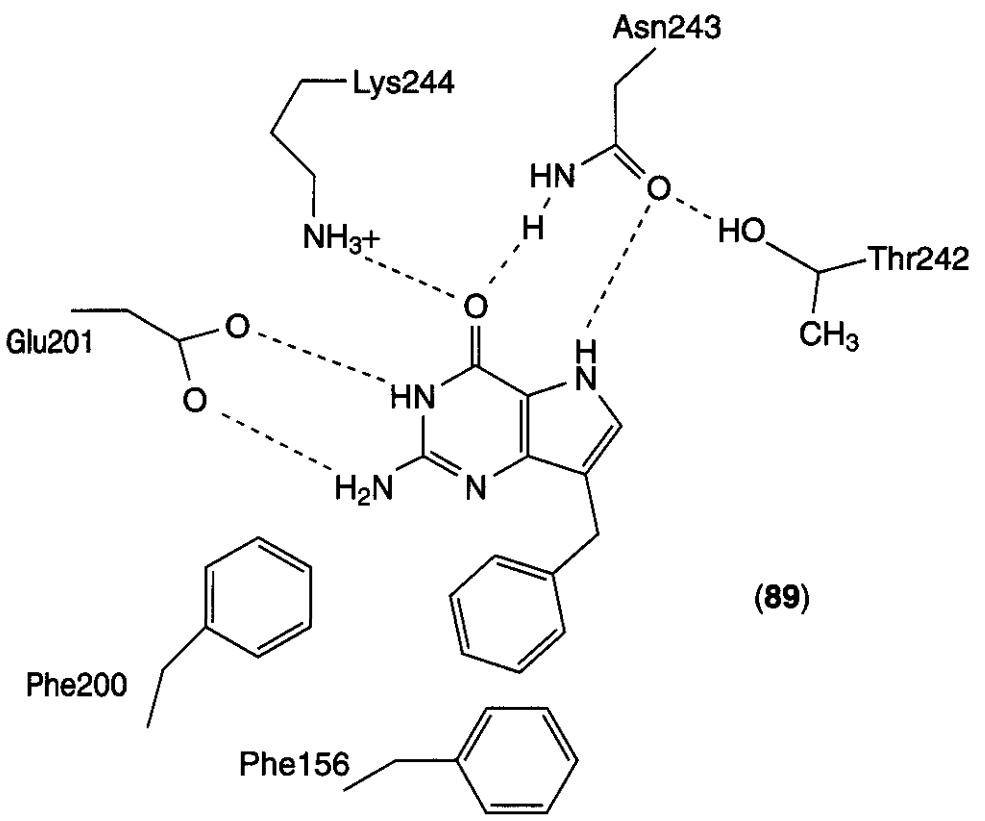
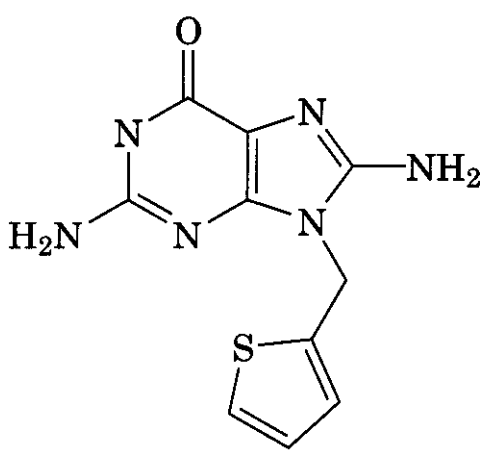


Figure 10.20. Binding interactions in the active site for the complex between PNP and (89).



(90)

other workers. This led to an exploration of a series of 8-mino, 9-deaza derivatives, although the hydrogen bonding for the simpler 9-deaza compounds turned out to be superior (reviewed in Ref. 202). It may be that compounds such as (90) suffer from unfavorable steric interactions between the 8-amino group and the proximal methyl group of Thr242 of the enzyme, or that the energetic cost of dehydrating the 8-amino group cannot be fully repaid by interactions with the enzyme. Other chemical series were also explored, but compound (86) had an acceptable safety profile and superior solubility and pharmacokinetic properties, and so was advanced into human testing.

2.9 Conclusions and Lessons Learned

The projects in which SBDD has been applied to enable the discovery of new drugs and clinical candidates have provided significant lessons for future investigators. Some of these' lessons learned are summarized here. Much of the credit for the summary presented here belongs to Michael Varney of Agouron, who provided a copy of a presentation that he made in 1998 to a medicinal chemistry symposium concerning the lessons learned in 10 years of use of SBDD methods.

Experience Matters. In every aspect of SBDD, as in all technical fields, there is no substitute for experience. Given the variety of different techniques that must be incorporated, this means that experience from several different people will be needed for optimal function of a discovery project team. Essential expertise is needed in X-ray crystallographic studies, graphical display of experimental results, initial and iterative design of compounds and synthetic tactics, the creation of databases and database queries, and the analyses of search outputs and of the results of computational simulation experiments.

Combine and Integrate Technologies. Dedicated molecular biology and protein chemistry personnel and equipment are essential for identifying the right constructs for crystallization and to the assurance of a steady supply of protein. Synthetic chemists trained in graphical analysis of protein structures tend to be excellent designers, and will be unlikely to design molecules that they cannot make. Early tactical integration of the synthetic approaches is even more important if combinatorial chemistry is part of the program. The structural information can be used to design combinatorial libraries as effectively as it can to design molecules one at a time. The use of libraries can compensate for the inaccuracies inherent in current computational scoring algorithms. More significantly, the integration of orthogonal technologies will stimulate creative thought and yield much more than the sum of the different technologies applied separately.

Go Big Early and Often. Filling active site space as much as possible will maximize the chance that a compound will be a potent inhibitor. During compound design, it should also be recognized that proteins are flexible, and that accessible conformations are hard to predict. Sometimes, larger functionality can be accommodated than the existing structural model permits. A few compounds should be included to probe this. These may give rise to an unexpected boon, such as access to a significantly altered new protein conformation with novel sites that can be exploited in new rounds of design and synthesis.

Aqueous Solubility is Critical to Success. Both early in SBDD and later on in clinical development, sufficient aqueous solubility is critical. Solubility is important early because the concentrations of compounds must be high during crystallization experiments to saturate the high levels of protein. The ratio of the solubility to the inhibition constant of a compound is also critical to the success of the crystallization experiment. Once some structural information becomes available, both parameters can be manipulated, but usually, soluble inhibitors must be available before the availability of structural information. Solubility matters during animal testing and later in development because compounds with very low solubility have limited or variable bioavailability.

Binding Sites Can Be Filled Many Ways. More than one small molecule scaffold can provide the necessary and sufficient hydrophobic and polar complementarity to generate potent inhibition. Sometimes, there are many scaffolds that will work. However, the structures of complexes with all the different scaffolds will likely have common features that are distinct from the structure of the apo enzyme, attributed to large-scale conformational changes that occur upon binding any ligand. The most useful X-ray models to use for the design of new compounds will be those that already have some substrate or inhibitor bound. There are several ways to design these compounds: modification of existing inhibitors, ***de novo*** creation of novel inhibitors, or some combination of these methods.

Not All Inhibitors Are Drugs. Having the X-ray structure of the target protein, or even having used the solved structure to design a potent inhibitor, is only the beginning of solving the difficult problems of drug design. The use of structure to create potent inhibitors can certainly shorten the time to get compounds into human testing, but use of SBDD methods does not guarantee that a potent compound will become a drug. This is an old lesson, actually, but is forgotten at great cost.

Structure of Free Inhibitor Is Important. Desolvation of the free ligand and of the protein's active-site groups upon complex formation are both significant. Both enthalpic and entropic contributions to the binding energy must be considered. Particular attention should be paid to the advantage that can be gained from "preorganization" of the inhibitor before binding, that is, low energy conformers bind with greater apparent avidity.

Bound Water Is Special, But Not All Hydrogen Bonds Are Created Equal. Each of the tightly bound waters present in an X-ray structural model has a unique environment and a unique function. In some cases, liberation of a bound water molecule by displacing it with an inhibitor's functionality can greatly increase inhibitor affinity, although this is not globally applicable. The entropic advantage of releasing a bound water into bulk solvent does

not always exceed the enthalpic cost of the displacement. In many situations, the preferred solution will be to retain a water molecule and use it to maximize inhibitor binding. For example, a water molecule that donates two hydrogen bonds and accepts one cannot be isosterically replaced. Electrostatic interactions that are more complex than hydrogen bonds and simple ion pairs are very difficult to model, anticipate, and exploit in inhibitor design.

Retain Potency While Addressing Other Issues. Structural information can be very useful in designing compounds that are not part of a competitor's intellectual property, or that cannot be patented because of information in the public domain. Redesign of a compound that is not itself proprietary, by use of structural information obtained with that compound, can yield valuable new proprietary molecules. Structural information can also guide the modification of physicochemical, metabolic, or pharmacological properties or target selectivity without compromising the potency against the primary therapeutic target.

All Models Are Wrong; Some Are Useful. At present, it is impossible to calculate an accurate value for a binding constant on an absolute scale. However, accurately estimating the relative binding of a series of closely related compounds is possible, and is much more likely to be successful if X-ray structures of target complexes with some of the compounds are available. Thus, although there is much room for improvement, local computational models can sometimes be quite useful. Even in the absence of an experimentally determined X-ray structure of the target, a hypothetical model can be a powerful tool for the design of useful compounds (e.g., captopril and **TAK-147).**

Iterative SBDD Cycles Are Optimal. Small alterations in ligand structure often cause major changes in binding mode, protein conformation, or both. These changes can go undetected if the structural effects are not analyzed by X-ray analysis iteratively or too infrequently. This can yield confusing or misleading structure-activity relationships, leading to a waste of precious time. Moreover, changes in compound structure seldom affect only one variable, so multiple orthogonal methods should be used to assess the effects of changes. It is also important during the rational design process to include room for serendipity. Do not reject an idea for a new compound that seems to make intuitive sense based on a single crystal structure or computational calculation.

REFERENCES

1. D. J. Abraham, ***Intra-Sci. Chem. Rep.,*** 8, 4 (1974).
2. http://www.agouron.com/
3. http://www.stromix.com/
4. http://www.astex-technology.com
5. http://www.accelrys.com/consortia/htc/
6. P. J. Goodford, **J. *Med. Chem.,*** 27,557 (1984).
7. C. R. Beddell, Ed., ***The Design of Drugs to Macromolecular Targets,*** John Wiley & Sons, Chichester, UK, 1992.
8. J. Greer, J. W. Erickson, J. J. Baldwin, and M. D. Varney, **J. *Med. Chem.,*** 37, 1035–1054 (1994).
9. R. E. Babine and S. L. Bender, ***Chem. Rev.***, 97, 1359 (1997).
10. P. Veerapandian, Ed., ***Structure-Based Drug Design,*** Marcel Dekker, New York, 1997.
11. R. T. Borchardt, R. M. Freidinger, T. K. Sawyer, and P. L. Smith, Eds., ***Integration of Pharmaceutical Discovery and Development. Case Histories*** (Pharmaceutical Biotechnology, Band 11), Plenum Press, New York, 1998.
12. K. Gubernator and H.-J. Bohm, Eds., ***Structure-Based Ligand Design,*** Wiley-VCH, New York/Weinheim, 1998.
13. C. L. Nobbs, H. C. Watson, and J. C. Kendrew, ***Nature,*** 209,339 (1966).
14. M. F. Perutz, ***Nature,*** 228, 726 (1970).
15. R. C. Ladner, E. J. Heidner, and M. F. Perutz, ***J. Mol. Biol.,*** 114,385 (1977).
16. G. Fermi, M. F. Perutz, B. Shaanan, and R. Fourme, **J. *Mol. Biol.,*** 175, 159 (1984).
17. B. C. Wishner, K. B. Ward, E. E. Lattman, and W. E. Love, **J. *Mol. Biol.,*** 98, 179 (1975).
18. D. J. Harrington, K. Adachi, and W. E. Royer Jr., **J. *Mol. Biol.,*** 272, 398 (1997).
19. C. R. Beddell, P. J. Goodford, G. Kneen, R. D. White, S. Wilkinson, and R. Wootton, ***Br. J. Pharmacol.,*** 82,397 (1984).
20. M. Merrett, D. K. Stammers, R. D. White, R. Wootton, and G. Kneen, *Biochem.* **J.**, 239,387 (1986).

21. F. C. Wireko and D. J. Abraham, ***Proc. Natl. Acad. Sci. USA,*** 88,2209 (1991).
22. D. J. Abraham, A. S. Mehanna, F. C. Wireko, E. P. Orringer, J. Whitney, and R. P. Thomas, ***Blood,*** 77, 1334 (1991).
23. M. K. Safo, S. Nokuri, and D. J. Abraham, Unpublished results.
24. P. E. Kennedy, F. L. Williams, and D. J. Abraham, ***J. Med. Chem.,*** 27, 103 (1984).
25. D. J. Abraham, M. F. Perutz, and S. E. V. Phillips, ***Proc. Natl. Acad. Sci. USA,*** 80,324 (1983).
26. M. F. Perutz, G. Fermi, D. J. Abraham, C. Poyart, and E. Bursaux, ***J. Am. Chem. Soc.,*** 108, 1064 (1986).
27. E. P. Orringer, D. S. Blythe, J. A. Whitney, S. Brockenbrough, and D. J. Abraham, ***Am. J. Hematol.,*** 39, 39 (1992).
28. D. J. Abraham, A. S. Mehanna, F. Williams, E. J. Cragoe Jr., and O. W. Woltersdorf Jr., ***J. Med. Chem.,*** 32,2460 (1989).
29. D. J. Abraham, P. E. Kennedy, A. S. Mehanna, D. Patwa, and F. L. Williams, ***J. Med. Chem.,*** 27,967 (1984).
30. A. Arnone, ***Nature,*** 237, 146 (1972).
31. V. Richard, G. G. Dodson, and Y. Mauguen, ***J. Mol. Biol.,*** 233,270 (1993).
32. P. J. Goodford, J. St-Louis, and R. Wootton, ***Br. J. Pharmacol.,*** 68, 741 (1980).
33. C. R. Beddell, P. J. Goodford, F. E. Norrington, S. Wilkinson, and R. Wootton, ***Br. J. Pharmacol.,*** 57,201 (1976).
34. F. F. Brown and P. J. Goodford, ***Br. J. Pharmacol.,*** 60,337 (1977).
35. A. S. Mehanna and D. J. Abraham, ***Biochemistry,*** 29,3944 (1990).
36. M. F. Perutz and C. Poyart, ***Lancet,*** 2, 881 (1983).
37. I. Lalezari and P. Lalezari, ***J. Med. Chem.,*** 32, 2352 (1989).
38. I. Lalezari, P. Lalezari, C. Poyart, M. Marden, J. Kister, B. Bohn, G. Fermi, and M. F. Perutz, ***Biochemistry,*** 29, 1515 (1990).
39. D. J. Abraham, R. S. Randad, M. A. Mahran, and A. S. Mehanna, ***J. Med. Chem.,*** 34, 752 (1991).
40. D. J. Abraham, F. C. Wireko, R. S. Randad, C. Poyart, J. Kister, B. Bohn, J. F. Leard, and M. P. Kunert, ***Biochemistry,*** 31,9141 (1992).
41. F. C. Wireko, G. E. Kellogg, and D. J. Abraham, ***J. Med. Chem.,*** 34,758 (1991).
42. D. J. Abraham, J. Kister, G. S. Joshi, M. C. Marden, and C. Poyart, ***J. Mol. Biol.,*** 248,845 (1995).
43. M. K. Safo, C. M. Moure, J. C. Burnett, G. S. Joshi, and D. J. Abraham, ***Protein Sci.,*** 10,951 (2001).
44. S. K. Burley and G. A. Petsko, ***FEBS Lett.,*** 201, 751 (1986).
45. S. K. Burley and G. A. Petsko, ***Science,*** **229**, 23 (1985).
46. M. Levitt and M. F. Perutz, ***J. Mol. Biol.,*** 201, 751 (1988).
47. D. J. Abraham, M. K. Safo, T. Boyiri, R. E. Danso-Danquah, J. Kister, and C. Poyart, ***Biochemistry,*** 34,15006 (1995).
48. M. P. Grella, R. Danso-Danquah, M. K. Safo, G. S. Joshi, J. Kister, S. J. Hoffman, M. Marden, and D. J. Abraham, ***J. Med. Chem.,*** 25, 4726 (2001).
49. A.M. Youssef, M. K. Safo, R. Danso-Danquah, G. S. Joshi, J. Kister, M. Marden, and D. J. Abraham, ***J. Med. Chem.,*** 45,1184 (2002).
50. J. A. Walder, R. H. Zaugg, R. Y. Walder, J. M. Steele, and I. M. Klotz, ***Biochemistry,*** 18,4265 (1979).
51. R. Chatterjee, E. V. Welty, R. Y. Walder, S. L. Pruitt, P. H. Rogers, A. Arnone, and J. A. Walder, ***J. Biol. Chem.,*** 261,9929 (1986).
52. S. R. Snyder, E. V. Welty, R. Y. Walder, L. A. Williams, and J. A. Walder, ***Proc. Natl. Acad. Sci. USA,*** 84,7280 (1987).
53. N. Komiyama, J. Tame, and K. Nagai, ***Biol. Chem.,*** 377,543 (1996).
54. T. Boyiri, M. K. Safo, R. E. Danso-Danquah, J. Kister, C. Poyart, and D. J. Abraham, ***Biochemistry,*** 34,15021 (1995).
55. M. F. Perutz, ***Br. Med. Bull.,*** 32, 195 (1976).
56. J. Monod, J. Wyman, and J.-P. Changeux, ***J. Mol. Biol.,*** **12**, 88 (1965).
57. D. A. Matthews, R. A. Alden, J. T. Bolin, S. T. Freer, R. Hamlin, N. Xuong, J. Kraut, M. Poe, M. Williams, and K. Hoogsteen, ***Science,*** 197, 452 (1977).
58. L. F. Kuyper, B. Roth, D. P. Baccanari, R. Ferone, C. R. Beddell, J. N. Champness, D. K. Stammers, J. G. Dann, F. E. Norrington, D. J. Baker, and P. J. Goodford, ***J. Med. Chem.,*** 25, 1120 (1982).
59. D. A. Matthews, J. T. Bolin, J. M. Burridge, D. J. Filman, K. W. Volz, and J. Kraut, ***J. Biol. Chem.,*** 260,392 (1985).
60. K. Appelt, R. J. Bacquet, C. A. Bartlett, C. L. J. Booth, S. T. Freer, M. A. Fuhry, M. R. Gehring, S. M. Herrmann, E. F. Howland, C. A. Janson, T. R. Jones, C. C. Kan, V. Kathardekar, K. K. Lewis, G. P. Marzoni, D. A. Matthews, C.

Mohr, E. W. Moomaw, C. A. Morse, S. J. Oatley, R. C. Ogden, M. R. Reddy, S. H. Reich, W. S. Schoettlin, W. W. Smith, M. D. Varney, J. E. Villafranca, R. W. Ward, S. Webber, S. E. Webber, K. M. Welsh, and J. White, J. Med. Chem., **34, 1925 (1991).**

61. S. H. Reich and S. E. Webber, Perspect. Drug Discov. Des., **1, 371-390 (1993).**

62. L. W. Hardy, J. S. Finer-Moore, W. R. Montfort, M. O. Jones, D. V. Santi, and R. M. Stroud, Science, **235, 448-455 (1987).**

63. D. A. Matthews, K. Appelt, S. J. Oakley, and N. H. Xuong, J. Mol. Biol., **214, 923-936 (1990).**

64. W. R. Montfort, K. M. Perry, E. B. Fauman, J. S. Finer-Moore, G. F. Maley, L. Hardy, F. Maley, and R. M. Stroud, Biochemistry, **29, 6964-6977 (1990).**

65. Y. Takemura and A. L. Jackman, Anticancer Drugs, **8, 3–16 (1997).**

66. S. E. Webber, T. M. Bleckrnan, J. Attard, J. G. Deal, V. Kathardekar, K. M. Welsh, S. Webber, C. A. Janson, D. A. Matthews, W. W. Smith, S. T. Freer, S. R. Jordan, R. J. Bacquet, E. F. Howland, C. L. J. Booth, R. W. Ward, S. M. Herrmann, J. White, C. A. Morse, J. A. Hilliard, and C. A. Bartlett, J. Med. Chem., **36, 733-746 (1993).**

67. I. Niculescu-Duvaz, Curr. Opin. Invest. Drugs, **2, 693-705 (2001).**

68. P. J. Goodford, J. Med. Chem., **28, 849 (1985).**

69. P. Goodford, J. Chemom., **10, 107 (1996).**

70. M. D. Varney, G. P. Marzoni, C. L. Palmer, J. G. Deal, S Webber, K. M. Welsh, R. J. Bacquet, C. A. Bartlett, C. A. Morse, C. L. Booth, S. M. Herrmann, E. F. Howland, R. W. Ward, and J. White, J. Med. Chem., **35, 663-676 (1992).**

71. D. R. Newell, Semin. Oncol., **26** (Suppl. 6), **74-81 (1999).**

72. P. Norman, *Curr.* Opin. Invest. Drugs, **2, 1611-1622 (2001).**

73. E. C. Taylor, Adv. Exp. Med. Biol., **338,** 387–**408 (1993).**

74. G. P. Beardsley, B. A. Moroson, E. C. Taylor, and R. G. Moran, J. Biol. Chem., **264, 328333 (1989).**

75. J. R. Piper, G. S. McCaleb, J. A. Montgomery, R. L. Kisliuk, Y. Gaumont, J. Thorndike, and F. M. Sirotnak, J. Med. Chem., **31, 2164-2169 (1988).**

76. S. E. Greasley, T. H. Marsilje, H. Cai, S. Baker, S. J. Benkovic, D. L. Boger, and I. A. Wilson, Biochemistry, **40, 13538-13547 (2001).**

77. R. J. Almassy, C. A. Janson, C. C. Kan, and Z. Hostomska, Proc. Natl. *Acad.* Sci. USA, **89, 6114-6118 (1992).**

78. C. C. Kan, M. R. Gehring, B. R. Nodes, C. A. Janson, R. J. Almassy, and Z. Hostomska, J. Protein Chem., **11, 467-473 (1992).**

79. M. D. Varney, C. L. Palmer, W. H. Romines 3rd, T. Boritzki, S. A. Margosiak, R. Almassy, C. A. Janson, C. Bartlett, E. J. Howland, and R. Ferre, *J.* Med. Chem., **40, 2502-2524 (1997).**

80. C. Shih, L. S. Gossett, J. F. Worzalla, S. M. Rinzel, G. B. Grindey, P. M. Harrington, and E. C. Taylor, J. Med. Chem., **35, 1109-1116 (1992).**

81. D. W. Cushman, H. S. Cheung, E. F. Sabo, and M. A. Ondetti, Biochemistry, **16, 5484** (1977).

82. M. A. Ondetti, B. Rubin, and D. W. Cushman, Science, **196, 441** (1977).

83. M. J. Wyvratt and A. A. Patchett, Med. Res. Rev., **5, 483-531 (1985).**

84. D. W. Cushman and M. A. Ondetti, Hypertension, **17, 589 (1991).**

85. D. W. Cushman and M. A. Ondetti, Nat. Med., **5, 1110 (1999).**

86. J. Rahuel, V. Rasetti, J. Maibaum, H. Rueger, R. Goschke, N. C. Cohen, S. Stutz, F. Cumin, W. Fuhrer, J. M. Wood, and M. G. Grutter, Chem. Biol., **7, 493-504 (2000).**

87. L. D. Byers and R. Wolfenden, Biochemistry, **12, 2070-2078 (1973).**

88. J. R. Huff and J. Kahn, Adv. Protein Chem., **56, 213-251 (2001).**

89. A. Wlodawer and J. Vondrasek, *Annu.* Rev. Biophys. Biomol. Struct., **27, 249** (1998).

90. T. D. Meek, *J.* Enzyme Inhib., **6,** 65 **(1992).**

91. A. Wlodawer and J. W. Erickson, *Annu.* Rev. Biochem., **62, 543 (1993).**

92. R. Lapatto, T. Blundell, A. Hemmings, J. Overington, A. Wilderspin, S. Wood, J. R. Merson, P. J. Whittle, D. E. Danley, K. F. Geoghegan, et al., Nature, **342, 299302 (1989).**

93. M. A. Navia, P. M. Fitzgerald, B. M. McKeever, C. T. Leu, J. C. Heimbach, W. K. Herber, I. S. Sigal, P. L. Darke, and J. P. Springer, Nature, **337, 615-620 (1989).**

94. A. Wlodawer, M. Miller, M. Jaskolski, B. K. Sathyanarayana, E. Baldwin, I. T. Weber, L. M. Selk, L. Clawson, J. Schneider, and S. B. Kent, Science, **245, 616-621 (1989).**

95. I. B. Duncan and S. Redshaw, Infect. Dis. Ther., **25,** 27–47 **(2002).**

96. A. G. Tomasselli, M. K. Olsen, J. O. Hui, D. J. Staples, T. K. Sawyer, R. L. Heinrikson, and C. S. Tomich, Biochemistry, **29, 264-269 (1990).**

97. M. Jaskolski, A. G. Tomasselli, T. K. Sawyer, D. G. Staples, R. L. Heinrikson, J. Schneider, S. B. Kent, and A. Wlodawer, ***Biochemistry,* 30, 1600–1609 (1991).**

98. M. W. Holladay, F. G. Salituro, and D. H. Rich, *J. Med. Chem.*, **30**, 374–383 (1987).

99. F. G. Salituro, N. Agarwal, T. Hofmann, and D. H. Rich, ***J. Med. Chem.*, 30, 286–295 (1987).**

100. D. J. Kempf, K. C. Marsh, D. A. Paul, M. F. Knigge, D. W. Norbeck, W. E. Kohlbrenner, L. Codacovi, S. Vasavanonda, P. Bryant, X. C. Wang, N. E. Wideburg, J. J. Clement, J. J. Plattner, and J. Erickson, ***Antimicrob. Agents Chemother.*, 35, 2209–2214 (1991).**

101. M. V. Hosur, N. T. Bhat, D. J. Kempf, E. T. Baldwin, B. Liu, S. Gulnik, N. E. Wideburg, D. W. Norbeck, K. Appelt, and J. W. Erickson, ***J. Am. Chem. Soc.*, 116, 847–855 (1994).**

102. D. J. Kempf, H. L. Sham, K. C. Marsh, C. A. Flentge, D. Betebenner, B. E. Green, E. McDonald, S. Vasavanonda, A. Saldivar, N. E. Wideburg, W. M. Kati, L. Ruiz, C. Zhao, L. Fino, J. Patterson, A. Molla, J. J. Plattner, and D. W. Norbeck, ***J. Med. Chem.*, 41, 602–617 (1998).**

103. C. N. Hodge, P. E. Aldrich, L. T. Bacheler, C. H. Chang, C. J. Eyermann, S. Garber, M. Grubb, D. A. Jackson, P. K. Jadhav, B. Korant, P. Y. Lam, M. B. Maurin, J. L. Meek, M. J. Otto, M. M. Rayner, C. Reid, T. R. Sharpe, L. Shum, D. L. Winslow, and S. Erickson-Viitanen, ***Chem. Biol.*, 3, 301-314 (1996).**

104. B. D. Dorsey, R. B. Levin, S. L. McDaniel, J. P. Vacca, J. P. Guare, P. L. Darke, J. A. Zugay, E. A. Emini, W. A. Schleif, J. C. Quintero, J. H. Lin, I. W. Chen, M. K. Holloway, P. M. D. Fitzgerald, M. G. Axel, D. Ostovic, P. S. Anderson, and J. R. Huff, ***J. Med. Chem.*, 37, 3443–3451 (1994).**

105. J. P. Vacca, J. P. Guare, S. J. DeSolms, W. M. Sanders, E. A. Giuliani, S. D. Young, P. L. Darke, I. S. Sigal, W. A. Schleif, J. C. Quintero, E. A. Emini, P. S. Anderson, and J. R. Huff, ***J. Med. Chem.*, 34, 1228–1230 (1991).**

106. T. A. Lyle, C. M. Wiscount, J. P. Guare, W. J. Thompson, P. S. Anderson, P. L. Darke, J. A. Zugay, E. A. Emini, W. A. Schleif, J. C. Quintero, R. A. F. Dixon, I. S. Sigal, and J. R. Huff, ***J. Med. Chem.*, 34, 1230–1233** (1991).

107. M. K. Holloway, J. M. Wai, T. A. Halgren, P. M. Fitzgerald, J. P. Vacca, B. D. Dorsey, R. B. Levin, W. J. Thompson, L. J. Chen, S. J. deSolms, N. Gaffin, A. K. Ghosh, E. A. Giuliani, S. L. Graham, J. P. Guare, R. W. Hungate, T. A. Lyle, W. M. Sanders, T. J. Tucker, M. Wiggins, C. M. Wiscount, O. W. Woltersdorf, S. D. Young, P. L. Darke, and J. A. Zuguay, ***J. Med. Chem.*, 38, 305–317 (1995).**

108. E. E. Kim, C. T. Baker, M. D. Dwyer, M. A. Murcko, B. G. Rao, R. D. Tung, and M. A. Navia, ***J. Am. Chem. Soc.*, 117, 1181–1182 (1995).**

109. S. W. Kaldor, V. J. Kalish, J. F. Davies 2nd, B. V. Shetty, J. E. Fritz, K. Appelt, J. A. Burgess, K. M. Campanale, N. Y. Chirgadze, D. K. Clawson, B. A. Dressman, S. D. Hatch, D. A. Khalil, M. B. Kosa, P. P. Lubbehusen, M. A. Muesing, A. K. Patick, S. H. Reich, K. S. Su, and J. H. Tatlock, ***J. Med. Chem.*, 40, 3979–3985 (1997).**

110. M. Moledina, M. Chakir, and P. J. Gandhi, *J. Thromb. Thrombolysis*, **12, 141–149 (2001).**

111. J. Hauptmann, ***Eur. J. Clin. Pharmacol.*, 57, 751–758 (2002).**

112. D. W. Banner and P. Hadvary, ***J. Biol. Chem.*, 266, 20085–20093 (1991).**

113. P. E. Sanderson and A. M. Naylor-Olsen, ***Curr. Med. Chem.*, 5, 289 (1998).**

114. J. P. Vacca, ***Curr. Opin. Chem. Biol.*, 4, 394** (2000).

115. J. Hauptmann, B. Kaiser, M. Paintz, and F. Markwardt, ***Biomed. Biochim. Acta*, 46, 445–453** (1987).

116. H. Brandstetter, D. Turk, H. W. Hoeffken, D. Grosse, J. Sturzebecher, P. D. Martin, B. F. Edwards, and W. Bode, ***J. Mol. Biol.*, 226, 1085–1099 (1992).**

117. N. H. Hauel, H. Nar, H. Priepke, U. Reis, J. M. Stassen, and W. Wienen, ***J. Med. Chem.*, 45, 1757–1766 (2002).**

118. R. M. Friedlander, V. Gagliardini, H. Hara, K. B. Fink, W. Li, G. MacDonald, M. C. Fishman, A. H. Greenberg, M. A. Moskowitz, and J. Yuan, ***J. Exp. Med.*, 185, 933–940 (1997).**

119. B. Siegmund, H. A. Lehr, G. Fantuzzi, and C. A. Dinarello, ***Proc. Natl. Acad. Sci. USA, 98, 13249–13254 (2001).**

120. N. P. Walker, R. V. Talanian, K. D. Brady, L. C. Dang, N. J. Bump, C. R. Ferenz, S. Franklin, T. Ghayur, M. C. Hackett, L. D. Hammill, L. Herzog, M. Hugunin, W. Houy, J. A. Mankovich, L. McGuiness, E. Orlewicz, M. Paskind, C. A. Pratt, P. Reis, A. Summani, M. Terranova, J. P. Welch, L. Xiong, A. Moller, D. E. Tracey, R. Kamen, and W. W. Wong, ***Cell*, 78, 343–352 (1994).**

121. K. P. Wilson, J. A. Black, J. A. Thomson, E. E. Kim, J. P. Griffith, M. A. Navia, M. A. Murcko, S. P. Chambers, R. A. Aldape, S. A. Raybuck, and D. Livingstone, ***Nature*, 370, 270–275 (1994).**

122. R. Leung-Toung, W. Li, T. F. Tam, and K. Karimian, *Curr. Med. Chem.*, 9,979–1002 (2002).

123. M. R. Michaelides and M. L. Curtin, *Curr. Pharm. Des.*, 5, 787–819 (1999).

124. P. D. Brown, *Expert Opin. Invest. Drugs*, 9, 2167–2177 (2000).

125. O. Santos, C. D. McDermott, R. G. Daniels, and K. Appelt, *Clin. Exp. Metastasis*, 15, 499–508 (1997).

126. L. J. MacPherson, E. K. Bayburt, M. P. Capparelli, B. J. Carroll, R. Goldstein, M. R. Justice, L. Zhu, S. Hu, R. A. Melton, L. Fryer, R. L. Goldberg, J. R. Doughty, S. Spirito, V. Blancuzzi, D. Wilson, E. M. O'Byrne, V. Ganu, and D. T. Parker, *J. Med. Chem.*, 40, 2525–2532 (1997).

127. G. Clemens, B. Hibner, R. Humphrey, H. Kluender, and S. Wilhelm in N. J. Clendeninn and K. Appelt, Eds., *Matrix Metalloproteinase Inhibitors in Cancer Therapy*, Humana Press, Totowa, NJ, 2001, pp. 175–192.

128. T. Kusumi, M. Tsuda, T. Katsunuma, and M. Yamamura, *Cell Biochem. Funct.*, 7, 201–204 (1989).

129. M. D. Sintchak, M. A. Fleming, O. Futer, S. A. Raybuck, S. P. Chambers, P. R. Caron, M. A. Murcko, and K. P. Wilson, *Cell*, 85, 921–930 (1996).

130. L. Hedstrom, *Curr. Med. Chem.*, 6, 545–560 (1999).

131. M. D. Sintchak and E. Nimmesgern, *Immunopharmacology*, 47, 163–184 (2000).

132. D. A. Gschwend, A. C. Good, and I. D. Kuntz, J. Mol. *Recognit.*, 9, 175–186 (1996).

133. D. Dvornik, *J. Diabetes Complications*, 6, 25–34 (1992).

134. D. R. Tomlinson, E. J. Stevens, and L. T. Diemel, *Trends Pharmacol. Sci.*, 15, 293–297 (1994).

135. C. L. Kaul and P. Ramarao, *Methods Find. Exp. Clin. Pharmacol.*, 23,465–475 (2001).

136. A. Urzhumtsev, F. Tete-Favier, A. Mitschler, J. Barbanton, P. Barth, L. Urzhumtseva, J. F. Biellmann, A. Podjarny, and D. Moras, *Structure*, 5,601–612 (1997).

137. M. C. Van Zandt, E. O. Sibley, K. J. Combs, E. E. McCann, B. Flam, D. J. Lavoie, D. Sawicki, A. Sabetta, A. Carrington, J. Sredy, V. Calderone, B. Cuevrier, and A. Podjarny, *Posterpresented at the 218th National Meeting of the American Chemical Society*, New Orleans, LA, August 22–26,1999.

138. S. Borman, *Chem. Eng. News*, 80, 35–39 (2002).

139. E. K. Perry, B. E. Tomlinson, G. Blessed, K. Bergrnann, P. H. Gibson, and R. H. Perry, *Br. Med. J.*,2,1457–1459 (1978).

140. B. P. Imbimbo, *CNS Drugs*, 15, 375–390 (2001).

141. Y. Ishihara, K. Kato, and G. Goto, *Chem. Pharm. Bull. (Tokyo)*, 39,3225–3235 (1991).

142. Y. Ishihara, G. Goto, and M. Miyamoto, *Curr. Med. Chem.*, 7, 341–354 (2000).

143. J. L. Sussman, M. Harel, F. Frolow, C. Oefner, A. Goldman, L. Toker, and I. Silman, *Science*, 253,872–879 (1991).

144. Y. Yamamoto, Y. Ishihara, and I. D. Kuntz, *J. Med. Chem.*, 37,314143153 (1994).

145. A. H. Reid, J. K. Taubenberger, and T. G. Fanning, *Microbes Infect.*, 3, 81–87 (2001).

146. J. N. Varghese, W. G. Laver, and P. M. Colman, *Nature*, 303, 35–40 (1983).

147. M. von Itzstein, W.-Y. Wu, G. B. Kok, M. S. Pegg, J. C. Dyason, B. Jin, T. Van Phan, M. L. Smythe, H. F. White, S. W. Oliver, P. M. Colman, J. N. Varghese, D. M. Ryan, J. M. Woods, R. C. Bethell, V. J. Hotham, J. M. Cameron, and C. R. Penn, *Nature*, 363,418–423 (1993).

148. P. Bossart-Whitaker, M. Carson, Y. S. Babu, C. D. Smith, W. G. Laver, and G. M. Air, *J. Mol. Biol.*, 232,1069–1083 (1993).

149. J. N. Varghese, V. C. Epa, and P. M. Colman, *Protein Sci.*, 4, 1081–1087 (1995).

150. C. U. Kim, W. Lew, M. Williams, H. Liu, L. Zhang, S. Swaminathan, N. Bischofberger, M. S. Chen, D. Mendel, W. G. Laver, and R. C. Stevens, *J. Am. Chem. Soc.*, 119,681 (1997).

151. Y. S. Babu, P. Chand, S. Bantia, P. Kotian, A. Dehghani, Y. El-Kattan, T. H. Lin, T. L. Hutchison, A. J. Elliott, C. D. Parker, S. L. Ananth, L. L. Horn, G. W. Laver, and J. A. Montgomery, *J. Med. Chem.*, 43, 3482–3486 (2000).

152. J. A. Green, G. M. Smith, R. Buchta, R. Lee, K. Y. Ho, I. A. Rajkovic, and K. F. Scott, *Inflammation*, 15, 355–367 (1991).

153. P. Vadas, J. Browning, J. Edelson, and W. Pruzanski, *J. Lipid Mediat.*, 8, 1–30 (1993).

154. C. Bennion, S. Connolly, N. P. Gensmantel, C. Hallam, C. G. Jackson, W. U. Primrose, G. C. Roberts, D. H. Robinson, and P. K. Slaich, *J. Med. Chem.*, 35,2939–2951 (1992).

155. S. Connolly, C. Bennion, S. Botterell, P. J. Croshaw, C. Hallam, K. Hardy, P. Hartopp, C. G. Jackson, S. J. King, L. Lawrence, A. Mete, D. Murray, D. H. Robinson, G. M. Smith, L. Stein, I. Walters, E. Wells, and W. J. Withnall, *J. Med. Chem.*, 45,1348–1362 (2002).

156. H. G. Beaton, C. Bennion, S. Connolly, A. R. Cook, N. P. Gensmantel, C. Hallam, K. Hardy, B. Hitchin, C. G. Jackson, and D. H. Robinson, J. Med. Chem., **37,557–559 (1994).**

157. J. P. Wery, R. W. Schevitz, D. K. Clawson, J. L. Bobbitt, E. R. Dow, G. Gamboa, T. Goodson Jr., R. B. Hermann, R. M. Kramer, D. B. McClure, et al., Nature, **352, 79–82 (1991).**

158. R. W. Schevitz, N. J. Bach, D. G. Carlson, N. Y. Chirgadze, D. K. Clawson, R. D. Dillard, S. E. Draheim, L. W. Hartley, N. D. Jones, Mihelich, et al., *Nat.* Struct. Biol., **2, 458–465 (1995).**

159. D. L. Scott, S. P. White, J. L. Browning, J. J. Rosa, M. H. Gelb, and P. B. Sigler, Science, **254,1007–1010 (1991).**

160. M. M. Thunnissen, E. Ab, K. H. Kalk, J. Drenth, B. W. Dijkstra, O. P. Kuipers, R. Dijkman, G. H. de Haas, and H. M. Verheij, Nature, **347,689–691 (1990).**

161. S. E. Draheim, N. J. Bach, R. D. Dillard, D. R. Berry, D. G. Carlson, N. Y. Chirgadze, D. K. Clawson, L. W. Hartley, L. M. Johnson, N. D. Jones, E. R. McKinney, E. D. Mihelich, J. L. Olkowski, R. W. Schevitz, A. C. Smith, D. W. Snyder, C. D. Sommers, and J. P. Wery, J. Med. Chem., **39,5159–5175 (1996).**

162. D. W. Snyder, N. J. Bach, R. D. Dillard, S. E. Draheim, D. G. Carlson, N. Fox, N. W. Roehm, C. T. Armstrong, C. H. Chang, L. W. Hartley, L. M. Johnson, C. R. Roman, A. C. Smith, M. Song, and J. H. Fleisch, *J.* Pharmacol. Exp. Ther., **288, 1117–1124 (1999).**

163. D. M. Springer, *Curr.* Pharm. Des., **7,181–198 (2001).**

164. C. Savolainen, S. Blomqvist, M. N. Mulders, and T. Hovi, J. Gen. Virol., **83 (Pt 2), 333–340 (2002).**

165. M. G. Rossmann, Viral Immunol., **2, 143–161 (1989).**

166. M. G. Rossmann, E. Arnold, J. W. Erickson, E. A. Frankenberger, J. P. Griffith, H.-J. Hecht, J. E. Johnson, G. Kamer, M. Luo, A. G. Mosser, R. R. Rueckert, B. Sherry, and G. Vriend, Nature, **317, 145–153 (1985).**

167. G. D. Diana, M. A. McKinlay, M. J. Otto, V. Akullian, and C. Oglesby, J. Med. Chem., **28, 1906–1910 (1985).**

168. G. D. Diana, M. A. McKinlay, C. J. Brisson, E. S. Zalay, J. V. Miralles, and U. J. Salvador, J. Med. Chem., **28, 748–752 (1985).**

169. M. J. Otto, M. P. Fox, M. J. Fancher, M. F. Kuhrt, G. D. Diana, and M. A. McKinlay, *Antimicrob.* Agents Chemother., **27, 883–886 (1985).**

170. M. P. Fox, M. J. Otto, and M. A. McKinlay, Antimicrob. Agents Chemother., **30, 110–116** (1986).

171. B. Jubelt, A. K. Wilson, S. L. Ropka, P. L. Guidinger, and M. A. McKinlay, J. Infect. Dis., **159, 866–871 (1989).**

172. T. J. Smith, M. J. Kremer, M. Luo, G. Vriend, E. Arnold, G. Kamer, M. G. Rossmann, M. A. McKinlay, G. D. Diana, and M. J. Otto, Science, **233,1286–1293 (1986).**

173. D. C. Pevear, M. J. Fancher, P. J. Felock, M. G. Rossmann, M. S. Miller, G. D. Diana, A. M. Treasurywala, M. A. McKinlay, and F. J. Dutko, *J.* Virol., **63**, 2002–2007 **(1989).**

174. K. H. Kim, P. Willingmann, Z. X. Gong, M. J. Kremer, M. S. Chapman, I. Minor, M. A. Oliveira, M. G. Rossmann, K. Andries, G. D. Diana, F. J. Dutko, M. A. McKinlay, and D. C. Pevear, *J.* Mol. Biol., **230, 206–227 (1993).**

175. G. D. Diana, D. Cutcliffe, R. C. Oglesby, M. J. Otto, J. P. Mallamo, V. Akullian, and M. A. McKinlay, *J.* Med. Chem., **32,450–455 (1989).**

176. G. D. Diana and D. C. Pevear, Antiviral Chem. Chemother., **8,401 (2002).**

177. G. D. Diana, P. Rudewicz, D. C. Pevear, T. J. Nitz, S. C. Aldous, D. J. Aldous, D. T. Robinson, T. Draper, F. J. Dutko, C. Aldi, et al., J. Med. Chem., **38,1355–1371 (1995).**

178. J. M. Rogers, G. D. Diana, and M. A. McKinlay, *Adv.* Exp. Med. Biol., **458**, 69–76 **(1999).**

179. F. G. Hayden, T. Coats, K. Kim, H. A. Hassman, M. M. Blatter, B. Zhang, and S. Liu, ***Antiviral*** Ther., **7, 53–65 (2002).**

180. B. Dérijard, J. Raingeaud, T. Barrett, I.-H. Wu, J. Han, R. J. Ulevitch, and R. J. Davis, Science, **267,682–685 (1995).**

181. K. P. Wilson, P. G. McCaffrey, K. Hsiao, S. Pazhanisamy, V. Galullo, G. W. Bemis, M. J. Fitzgibbon, P. R. Caron, M. A. Murcko, and M. S. Su, Chem. Biol., **4,423–431 (1997).**

182. B. Frantz, T. Klatt, M. Pang, J. Parsons, A. Rolando, H. Williams, M. J. Tocci, S. J. O'Keefe, and E. A. O'Neill, Biochemistry, **37, 13846–13853 (1998).**

183. J. C. Lee, J. T. Laydon, P. C. McDonnell, T. F. Gallagher, S. Kumar, D. Green, D. McNulty, M. J. Blumenthal, J. R. Heys, S. W. Landvatter, J. E. Strickler, M. M. McLaughlin, I. R. Siemens, S. M. Fisher, G. P. Livi, J. R. White, J. L. Adams, and P. R. Young, Nature, **372, 739–746 (1994).**

184. J. C. Lee, S. Kumar, D. E. Griswold, D. C. Underwood, B. J. Votta, and J. L. Adams, *Immunopharmacology*, **47,185–201 (2000).**

185. A. Cuenda, J. Rouse, Y. N. Doza, R. Meier, P. Cohen, T. F. Gallagher, P. R. Young, and J. C. Lee, FEBS Lett., **364, 229-233(1995).**

186. A. M. Badger, J. N. Bradbeer, B. Votta, J. C. Lee, J. L. Adams, and D. E. Griswold, *J. Pharmacol. Exp. Ther.*, **279**, 1453–1461 (1996).

187. Z. Wang, B. J. Canagarajah, J. C. Boehm, S. Kassisa, M. H. Cobb, P. R. Young, S. Abdel-Meguid, J. L. Adams, and E. J. Goldsmith, Structure, **6, 1117-1128(1998).**

188. K. P. Wilson, M. J. Fitzgibbon, P. R. Caron, J. P. Griffith, W. Chen, P. G. McCaffrey, S. P. Chambers, and M. S. Su, J. Biol. Chem., **271, 27696-27700 (1996).**

189. T. Fox, J. T. Coll, X. Xie, P. J. Ford, U. A. Germann, M. D. Porter, S. Pazhanisamy, M. A. Fleming, V. Galullo, M. S. Su, and K. P. Wilson, Protein Sci., **7, 2249(1998).**

190. R. J. Gum, M. M. McLaughlin, S. Kumar, Z. Wang, M. J. Bower, J. C. Lee, J. L. Adams, G. P. Livi, E. J. Goldsmith, and P. R. Young, J. Biol. Chem., **273, 15605-15610(1998).**

191. J. L. Adams, J. C. Boehm, T. F. Gallagher, S. Kassis, E. F. Webb, R. Hall, M. Sorenson, R. Garigipati, D. E. Griswold, and J. C. Lee, Bioorg. Med. Chem. Lett., **11, 2867-2870 (2001).**

192. T. Fullerton, A. Sharma, U. Prabhakar, M. Tucci, S. Boike, H. Davis, D. Jorkasky, and W. Williams, Clin. Pharmacol. Ther., **67, 114 (2000).**

193. Pat. Appl. Vertex Pharmaceuticals, Inc., assignee, PCT WO 00/36096 **(2000).**

194. J. J. Haddad, *Curr.* Opin. Invest. Drugs, **2, 1070** (2001).

195. C. Pargellis, L. Tong, L. Churchill, P. F. Cirillo, T. Gilmore, A. G. Graham, P. M. Grob, E. R. Hickey, N. Moss, S. Pav, and J. Regan, *Nat.* Struct. Biol., **9, 268-272 (2002).**

196. J. Regan, S. Breitfelder, P. Cirillo, T. Gilmore, A. G. Graham, E. Hickey, B. Klaus, J. Madwed, M. Moriak, N. Moss, C. Pargellis, S. Pav, A. Proto, A. Swinamer, L. Tong, and C. Torcellini, J. Med. Chem., **45, 2994(2002).**

197. G. R. Boss and J. E. Seegmiller, *Annu.* Rev. Genet., **16**, 297–328 **(1982).**

198. S. E. Ealick, S. A. Rule, D. C. Carter, T. J. Greenhough, Y. S. Babu, W. J. Cook, J. Habash, J. R. Helliwell, J. D. Stoeckler, R. E. Parks Jr., S. Chen, and C. E. Bugg, J. Biol. Chem., **265, 1812(1990).**

199. S. E. Ealick, Y. S. Babu, C. E. Bugg, M. D. Erion, W. C. Guida, J. A. Montgomery, and J. A. Secrist 3rd, *Proc.* Natl. Acad. Sci. USA, **88, 11540-11544(1991).**

200. J. A. Montgomery, S. Niwas, J. D. Rose, J. A. Secrist 3rd, Y. S. Babu, C. E. Bugg, M. D. Erion, W. C. Guida, and S. E. Ealick, *J.* Med. Chem., **36**, 55–69 **(1993).**

201. M. Duvic, E. A. Olsen, G. A. Omura, J. C. Maize, E. C. Vonderheid, C. A. Elmets, J. L. Shupack, M. F. Demierre, T. M. Kuzel, and D. Y. Sanders, *J. Am.* Acad. Dermatol., 44, **940-947** (2001).

202. P. E. Morris Jr. and G. A. Omura, *Curr.* Pharm. Des., **6, 943-959 (2000).**

CHAPTER ELEVEN

X-Ray Crystallography in Drug Discovery

DOUGLAS A. LMNGSTON
SEAN G. BUCHANAN
KEVINL. D'AMICO
MICHAEL V. MILBURN
THOMAS S. PEAT
J. MICHAEL SAUDER
Structural GenomiX
San Diego, California

Contents

1 Introduction, **472**
2 Methodology, **472**
  2.1 Theory, **472**
  2.2 Crystallization, **473**
  2.3 Data Collection, **474**
  2.4 Phase Problem, **476**
  2.5 Computing and Refinement, **478**
  2.6 Databases, **478**
3 Applications of the Use of Crystallographic Studies in Drug Discovery and Development, **479**
4 Structural Genomics, **481**
  4.1 Introduction to Structural Genomics, **481**
  4.2 Genome Annotation, **481**
  4.3 Pathways, **495**
  4.4 Protein Structure Modeling, **495**
5 Conclusion, **496**

Burger's Medicinal Chemistry and Drug Discovery
Sixth Edition, Volume **1:** Drug Discovery
Edited by Donald **J.** Abraham
ISBN 0-471-27090-3 © **2003** John Wiley & Sons, Inc.

1 INTRODUCTION

The practice of crystallography is undergoing dramatic change because of the advent of new robotics technologies, orders-of-magnitude improvement in X-ray sources and computational power, and the advances in protein production stemming from the recent revolution in molecular biology. This chapter covers these changes in the context of an overview of the techniques of modern crystallography, their application in the identification and characterization of targets and mechanisms for therapeutic intervention, and the nascent field of structural genomics. Structure-based drug design applications are covered elsewhere.

The exponential growth in the rate of determination of new protein structures continues unabated. Technologies developed in the late 1980s (1) have now evolved to the point that they have been implemented in high-throughput (HTS) format, driving the rate even higher. Super-intense, precise, tunable X-rays are now available from undulator beamlines. Three "third-generation" synchrotrons, designed and built for this purpose, are now on line—ESRF in Grenoble, France; Spring-8 in Japan; APS at Argonne National Laboratory in the United States—and others are under construction. In a relative sense, this capability has had minimal impact on medicinal chemistry to date, but that will certainly change. The companies that have successfully built high-throughput protein crystallography systems (SGX and Syrrx in the United States and Astex in the U.K. among others) have all now turned their prodigious capacity to the co-crystallization of small molecules with target proteins for the purpose of drug discovery. The capacity to compare, in parallel, the binding modes of a set of hits from HTS, or a given lead series, will be valuable, but an even greater impact will result from the decrease in turnaround time required to generate co-crystal structures. This has been the most significant hindrance to realizing the full potential of structure-based drug design. A structure is far more useful before the chemist has embarked on the synthesis of the next series, rather than after.

Another important development toward new target identification is the effort in large-scale structural annotation of various genomes, the field of structural genomics. In classifying proteins by function as a step toward validating them as therapeutic targets, structural homology is perhaps the most important tool available. These efforts have been taken up by a number of publicly-funded consortia (2), because the commercial value of genomic databases in general has not been high enough to justify their cost in the private sector. Given that medicinal chemists think and communicate largely in structural terms, this recent growth in the influence of structural biology is very important. It forms the basis of a powerful link between chemistry and biology, and we have only begun to realize its potential.

2 METHODOLOGY

2.1 Theory

X-ray crystallography provides atomic or near atomic resolution of matter. The periodicity of crystals, reflecting the repeating units of molecular structure, diffracts X-rays according to Bragg's law: $n\lambda = 2d\sin\theta$, where n is the order of diffraction, λ the wavelength of the radiation, d the spacing or distance between a family of lattice planes in the crystal, and e the angle of the diffraction. X-radiation is ideal to analyze atomic structure, because the wavelengths used are in the order of 0.1–2.0 Å with 0.75 Å being about one-half the distance of an aliphatic carbon-carbon bond.

The images of diffracted crystal lattices can be observed with specialized precession photographic equipment, although the modern day image plate detectors used in most laboratories produce a diffraction image that can be analyzed by computer to provide the indices of the lattice diffraction spots (Fig. 11.1, a–c).

The X-ray diffraction from the electron clouds surrounding each nucleus is either reinforced or impeded and gives rise to the dif-

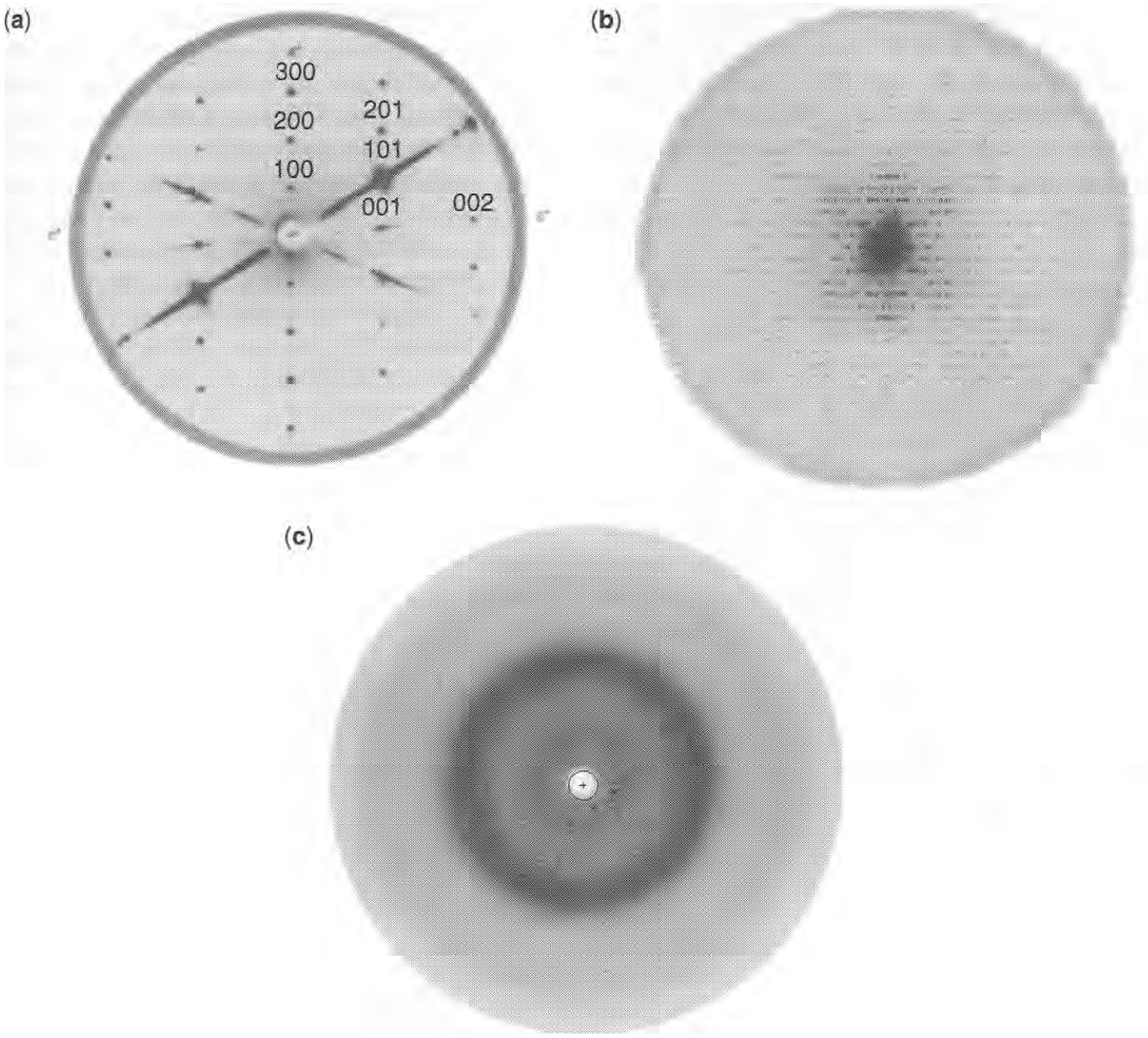


Figure **11.1.** (a) A look at a two-dimensional crystal lattice diffraction pattern for a small molecule natural product, MW 222. Each diffraction intensity in the lattice is numbered to give a unique three dimensional address (identification)for that measurement. These numerical addresses are referred to as Miller indices or hkl values. (b) A diffraction pattern from a precession photograph for hemoglobin, MW 65,000. Note the the diffraction lattice spacings are much smaller for the large molecule and reflects the nature of Bragg's law, where the lattice is observed in reciprocal space ($1/d = 2\sin\theta/n\lambda$). (c) **An** image plate diffraction pattern for a protein. [Adapted with permission from D. J. Abraham, Computer-Aided Drug Design, Methods, and Applications, Marcel Dekker, Inc., New York, 1989.1

ference in intensities observed in Fig. 11.1. The steps that one goes through to solve a crystal structure follow, with the intent of providing the non-crystallographer with a simplified and pictorial view of the process.

2.2 Crystallization

Crystallization is the critical first and most important step, because good single crystals usually provide quality diffraction. Linus Paulling once entitled one of his lectures "The Importance of Being Crystalline" **(3).** Unfortunately, crystallization is still more empirical than scientific. It requires closely monitored matrix changes in growing conditions, i.e., pH, salt concentration, temperature, solvents, and crystallization setups. Most laboratories now use well-known sparse matrix screens pioneered by Jancarik and Kim (4) and further refined and commercially distributed by Hampton Research **(5,** 6). Screens will typically employ vapor diffusion experiments (hangingdrops or sitting drops), and occasionally batch and liquid-liquid diffusion methods.

More recently, batch crystallizations have been rejuvenated by the development of microbatch robots and by the groups of Chayen (7), DeTitta (8), and D'Arcy (9).

Although discovering the crystallization conditions for a new protein or nucleic acid can be tedious, relatively inexperienced individuals can usually succeed at growing crystals once the initial conditions are established. Some of the most successful crystallization methodologies are based on vapor diffusion methods (Fig. 11.2). The general idea behind vapor diffusion crystallization is to dissolve the protein in a buffer, with a non-precipitating amount of the miscible vapor solvent, in a reservoir that is in equilibrium with higher concentration of the vaporizing solvent nearby. Another variation is to set up the crystallization cocktail containing salts, buffers, P (poly) ethylene glycols (PEGS), small molecule solvents, etc., where volume is slowly reduced by the equilibrating mixture, which is placed nearby. McPherson, Carter, and others have developed more quantitative methods for optimizing crystal growth (10).

2.3 Data Collection

Most laboratories have rotating anode sources for production of high intensity X-ray beams. These are coupled with an area detector that has made single crystal diffractometers obsolete. Mirrors and other technology have also been used to provide a more intense and monochromatic radiation source (11).Radiation from rotating anode sources is at a fixed wavelength, usually from high-voltage electrons impinging on either a copper or molybdenum rotating anode, i.e., radiation at 1.54 Å (copper) or 0.71 Å (molybdenum). Radiation from synchrotron sources can often be tuned to a wavelength of interest for multiwavelength anomalous diffraction (MAD) experiments (see below).

X-rays generated by a synchrotron source are typically two orders of magnitude stronger than conventional CuKα radiation generated by a rotating anode. Synchrotron sources have greatly extended the ability to solve new protein structures when only weakly diffracting or small crystals are available. Another advantage in using the stronger synchrotron radiation is that the crystal exposure time is significantly lower. The typical exposure time for home laboratory CuKa sources ranges from 5 to 60 min for a range of data, whereas the equivalent set of data at an undulator beamline, i.e., the advanced photon source (APS), requires only about 1 s of exposure time. Synchrotron radiation has also allowed the use of MAD, enabling phasing (imaging) of the protein using a derivative with only one heavy element.

A variety of detectors are in common use to record X-ray data and have the advantage of measuring the intensities of large numbers of diffraction spots simultaneously. The most popular detectors are image plates and charge-coupled device (CCD) cameras. Image plates are typically the choice for laboratory rotating anode sources and lower flux synchrotron sources (Fig. 11.3). CCDs have the distinct advantage of speed at the higher flux synchrotron sources, because they simultaneously measure and record diffraction intensities (amplitudes). Current CCD cameras have readout times on the order of a few (typically 2–8) seconds, a speed not dreamed of when the first protein structure data was recorded from photographs (with intensities measured by eye comparison to standard reference spots on a separate film strip). Speed of data collection can be an important advantage at third generation synchrotron sources, with even shorter exposure times. On the other hand, image plates have a greater range of use, being accessible in any X-ray diffraction laboratory, with many of the newer models taking less than 1 min to record the intensity data. Image plate detectors often have more than one image plate, so one can be read while the other is exposed, effectively wasting no time during the collection period. The image plates also offer a larger surface area for data collection than most CCD cameras and are considerably less expensive.

X-ray diffraction data from crystals are either collected at room temperature or under cryogenic conditions at liquid nitrogen temperatures [around 100°K (–170°C)]. For room temperature data collection, crystals are normally mounted in thin-walled glass capillaries, with a small amount of mother liquor about 5 mm from the crystal. The mother liquor in the capillary is critical because protein crystals are 40–80% water—dried protein crystals do not diffract. The nearby mother

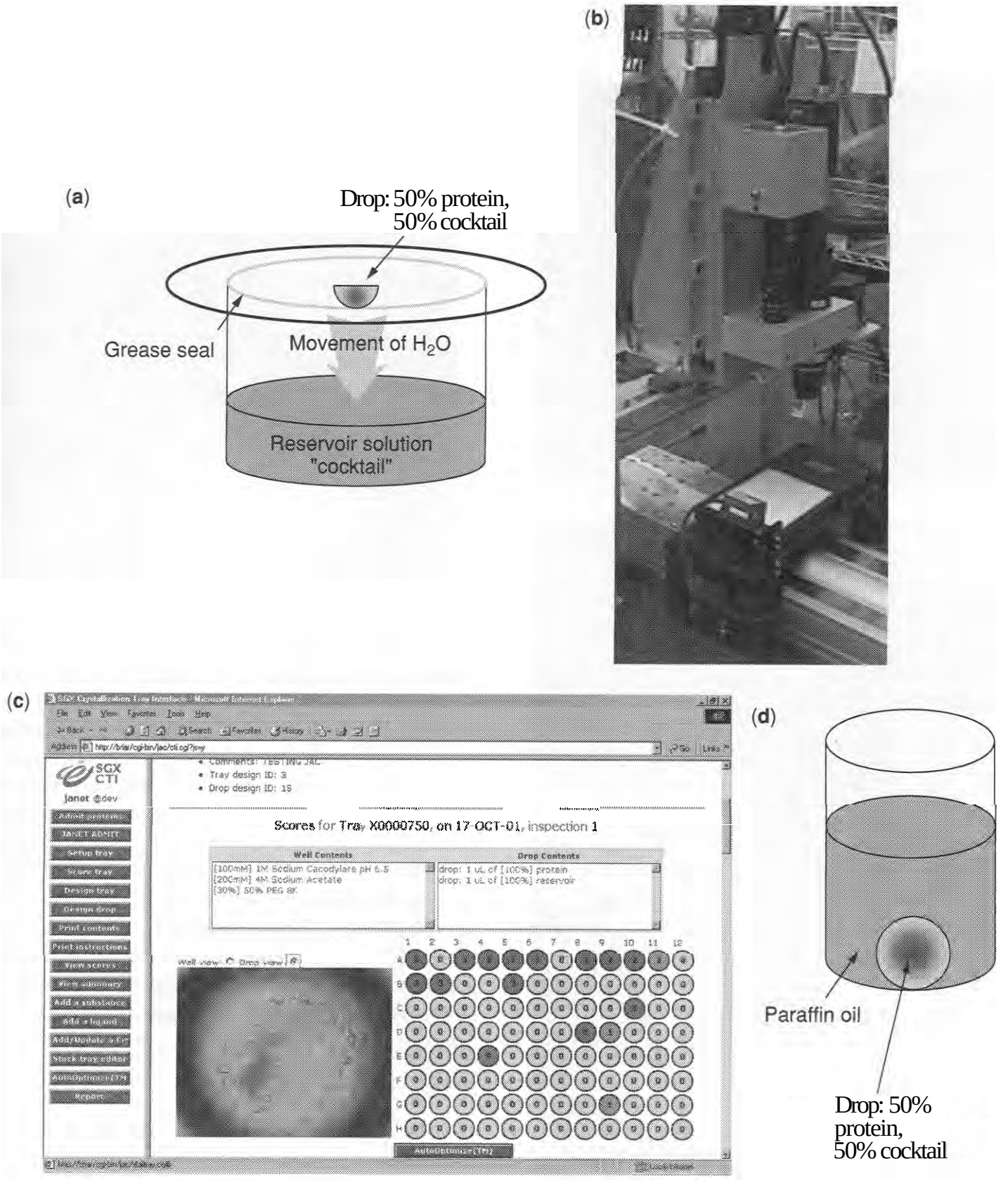


Figure 11.2. (a) The drops are typically 1–10 $\mu\lambda$ total volume, with between 100 and 1000 $\mu\lambda$ total volume of cocktail in the well. The smaller the drop size, the faster the equilibrium occurs, in general. There are a variety of plates now available in which to set up these vapor diffusion experiments, the most common being 24-well Limbro plates and 96-well microtiter plates. Several robots have been developed to automatically set up the crystallization experiments; although most are no faster than doing the same procedure by hand (particularly with a multi-channel pipettor), there can be other advantages (e.g., consistency and reducing repetitive stress syndrome). Once plates are set up, they are typically kept at a constant temperature and observed periodically under a microscope. (b and c) Progress in automating this aspect of characterization has occurred, and there are now imagers that will take high resolution, digital pictures of each drop in turn and store these for either manual or automated analysis. (d) Batch experiments are set up such that the protein is mixed with cocktail and there is little concentration or dilution to the sample over time. This can now be done in very high throughput and small scale: 50–200 nL drops under oil in 1536-well plates, for example. This kind of approach has been used to screen hundreds of conditions with small amounts of protein, which may allow for faster optimization later. One caveat is that small crystals don't necessarily lead to larger crystals later, and all structures to date have had crystals of greater than 10 microns in at least one dimension.

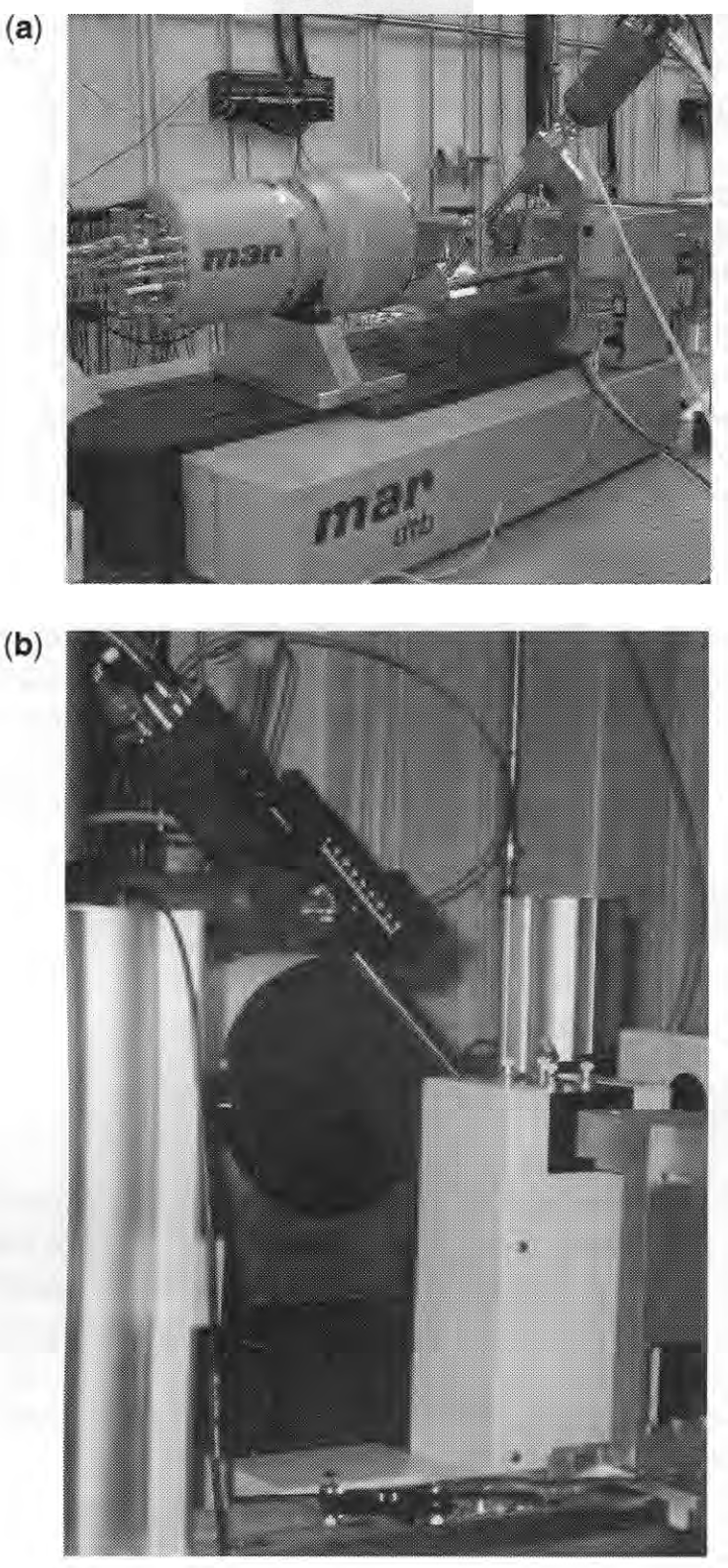


Figure 11.3. (a) Area detector showing the configuration of the unit. (b) Area detector showing the face.

liquor ensures that the crystal is bathed in the vapor of the mother liquor and prevents drying during data collection. The majority of present day data collections in home laboratories and at synchrotrons are done under cryogenic conditions, which allows high intensity X-radiation to be used without the crystal decay observed in room temperature data collection. For cryogenic data collection, crystals are normally mounted in a thin fiber loop with a layer of suitable cryoprotectant solution (Fig. 11.4). The cryoprotectant forms a layer of noncrystalline glass around the crystal to protect it from freeze shock. Simple freezing of the crystal results in the formation of ice in the interior of the crystal and renders it useless. A quick perusal of the literature shows PEG, glycerol, sucrose, and 2-methyl-2,4-pentane diol (MPD) as the most popular cryoprotectants. Oils, such as paraffin oil, have also been used successfully as cryoprotectants (12).

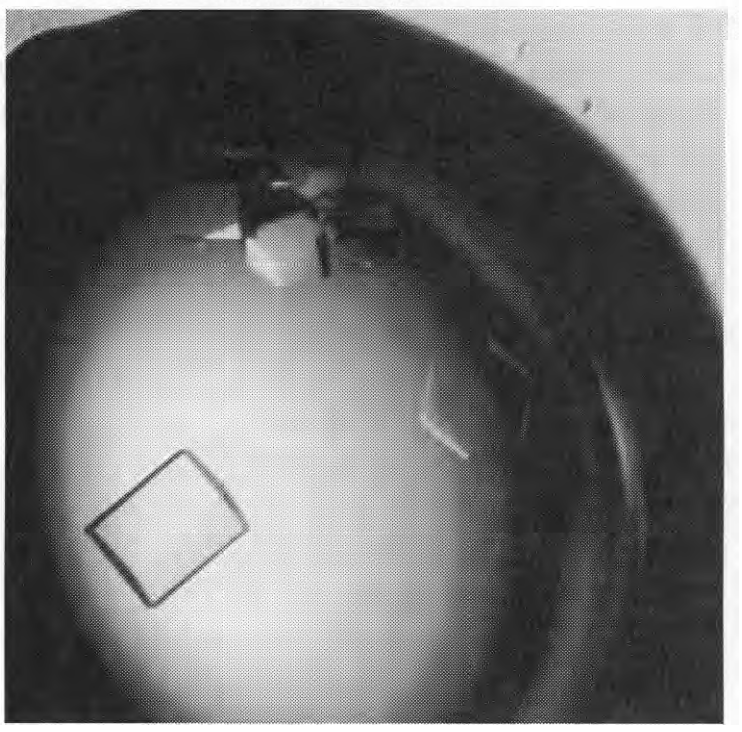


Figure 11.4. The crystals are manipulated by scooping them up with a small loop of nylon that is glued to the end of a pin. Surface tension from the liquid will hold the crystal in the loop, but the crystal can also be held by using a loop that is smaller in size than the crystal of interest. This technique will work particularly well with fragile crystals, thin plates for example, that would normally fall apart in a capillary mount. Once the crystal is frozen, it is placed on an axis in line of both an X-ray source and a stream of nitrogen set to about 100,000 to keep the crystal frozen. The crystal is rotated in increments during the data collection procedure to collect a full data set (typically one or two degrees per frame, depending on the resolution limits, mosaicity of the crystal, unit cell lengths, etc.).

2.4 Phase Problem

X-ray diffraction measurements as described above only provide the amplitudes of the diffracted waves. One must have the phase angles of all measured waves relative to a common origin in order to image the molecule using a Fourier analysis. Figure 11.5 illus-

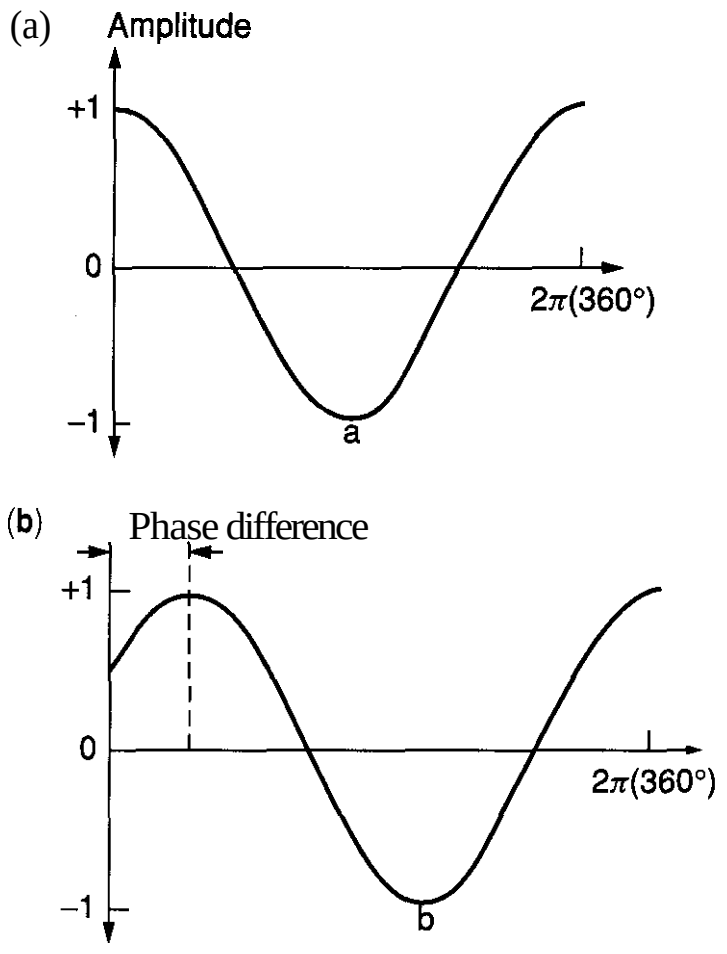


Figure 11.5. Graphs showing the phase relationship of electromagnetic radiation. [Adapted with permission from D. J. Abraham, Computer-Aided Drug Design, Methods, and Applications, Marcel Dekker, Inc., New York, 1989.1

trates the differences in the phases and amplitudes for two reflections. The solution of the phase problem that permitted the first image reconstruction of a protein was discovered by Perutz using multiple isomorphous replacement (MIR) **(13).**

The majority of the earliest structures were solved using MIR to phase the maps. This requires soaking the crystals or co-crystallizing the protein with two or more heavy atoms and hoping that these heavy atoms bind in a specific way to the protein. It also requires that the subsequent crystals are isomorphous with the native protein (i.e., no changes in the unit cell or symmetry of the crystal). Although it is possible to obtain phase information from a single heavy atom derivative using additional information (e.g., anomalous scattering or density modification), one often works diligently to get a second or third derivative to improve the quality of the electron density (Fourier) map.

Two other common methods are used to estimate phases in protein crystallography: molecular replacement (MR), which uses the structural motif of **a** homologous protein (14), and MAD from **a** single heavy element (1).

MR methodology requires a structural model that is structurally homologous to the protein that has been crystallized. Phasing is accomplished through a six-dimensional search—a three-dimensional rotation search followed by a three-dimensional translation search, using the model against the crystal data. Molecular replacement is being employed more frequently as the number of known structures has increased, which has made unique structural motifs available for phasing. For highly homologous protein structures, this method is usually straightforward and successful. For marginal cases, the addition of some independent phase information, single isomorphous replacement (SIR) or MIR, in combination with MR can enhance the quality of the Fourier map.

MAD phasing is an alternate methodology for solving the phase problem. MAD requires a single heavy atom with anomalous peak scattering at **a** wavelength where X-rays both at and near the spectral energy are accessible. Data sets are collected at different wavelengths to optimize the anomalous and dispersive signals from the heavy atom. Certain beamlines have been designed with wavelengths that are tunable "on the fly," and these are often referred to as MAD beamlines. MAD has become the method of choice for rapid structure solution when synchrotron radiation is available. The advantage of MAD phasing is that one often only needs a single crystal to collect all of the data necessary to solve the structure. Although multiple wavelengths are collected (anywhere from two to five sets), data collection is routinely completed in less than a few hours. The peak wavelength choice data set is very important to collect first as it contains the greatest anomalous signal and is often used alone to find the heavy atom sites (Shake'n Bake program, anomalous Patterson maps, etc.). If the crystal degrades quickly in the beam, one can also employ single-wavelength anomalous diffraction phasing (SAD) if a full data set at the peak wavelength was successfully collected. SAD phasing requires additional information, obtained by density modification, to obtain interpretable electron density maps, but has been

proven in many instances to result in very high quality maps (15).

Many different heavy atoms have been used for MAD/SAD phasing, the most popular being selenium. Selenium is incorporated into the amino acid sequence of the protein by adding selenomethionine to the growth media when the protein is produced (16). For proteins that bind DNA, 5-bromouracil has been a popular choice for phasing through anomalous scattering. Most heavy elements have good anomalous signals (Hg, Pt, U, Au, etc.). Lanthanides have a particularly good signal and can sometimes substitute for divalent metals found naturally in the protein (e.g., Ca) (17). One of the major advantages of MAD phasing is that the signal does not decay at higher resolution with perfectly isomorphic crystals, so the experimentally phased map can be quite good out to the full resolution of diffraction. This typically has not been the case when using multiple isomorphous replacement, where the experimentally phased map often only extends to around 2.5 Å resolution, because of a lack of isomorphism between the native and heavy metal substituted crystals. Anomalous scattering has been useful in the structure determination of very large structures; the 30S ribosome was recently solved using Os and Lu derivatives (18).

2.5 Computing and Refinement

Raw intensity X-ray crystallographic data is next reduced and scaled to provide structure factors (F) that are used to solve and image the structure. Two of the most popular software packages employed to reduce raw date are Mosflm/CCP4 (19) and the HKL suite (20). Both work very well and are very fast with modern computers. A variety of programs, such as SOLVE (21), Shake and Bake (22), or SHELX (23), can be employed to find the heavy atom positions, including hand searching methods through Patterson maps. Once heavy atom sites are found, they are usually refined with the programs SHARP (24) or MLPHARE (19). The heavy atom positions are next used as phase information input to provide initial phases for electron density maps, which are used to fit the remainder of the protein or nucleic acid. Once a model of the structure is obtained it is refined. In cases where high resolution data is available, programs such as wARP (24) can automatically provide models of protein structures. When high resolution data is not available. a model is most often built in by hand using such graphics programs like O (25) or XFIT. The models are refined against the data by programs such as REFMAC (19) and CNS (26). All of these programs have become much faster and easier to use because of the incredible increases in speed that new hardware has allowed.

It is worth mentioning that statistical and probabilistic techniques have had a significant impact in how heavy atoms are found and models are refined (e.g., SHARP, SOLVE, REFMAC). Baysian statistics and maximum likelihood methods are now used instead of least-squares methods. One may want to consider how various data collection strategies may affect the later steps in the process by keeping this in mind, i.e., high redundancy in the data makes for better statistics.

The quality of a structure is measured in many ways: how low the ***R*** factor or R_{free} is (the fit of observed data to the model), the resolution limit of the data, the ideality of the bonds and angles, etc. How well a structure measures up to other structures of about the same resolution also gives a good idea of how "good" a given structure is (PROCHECK program). SFCHECK is a useful program for assessing the agreement between the atomic model and the experimental X-ray data. The level of confidence one expects from a given model will depend on the resolution of the data. This can be seen clearly in Fig. 11.6, where a residue from a protein structure is shown with three different data cutoffs at different resolution ranges. The model from a 3.0-Å data set may look the same as one from a 1.3-Å data set, but the level of confidence is much higher in the latter. A reasonably well-refined structure will have a crystallographic ***R*** factor between 15% and 25% and will have an R_{free} of less than 30% under most circumstances.

2.6 Databases

The Protein Data Bank (PDB) (27, 28) is now coordinated by a consortium of several institutions (Rutgers University, the San Diego Supercomputer Center, and National Institute for Standards and Technology). As of this writing, the PDB has over 18,000 structures,

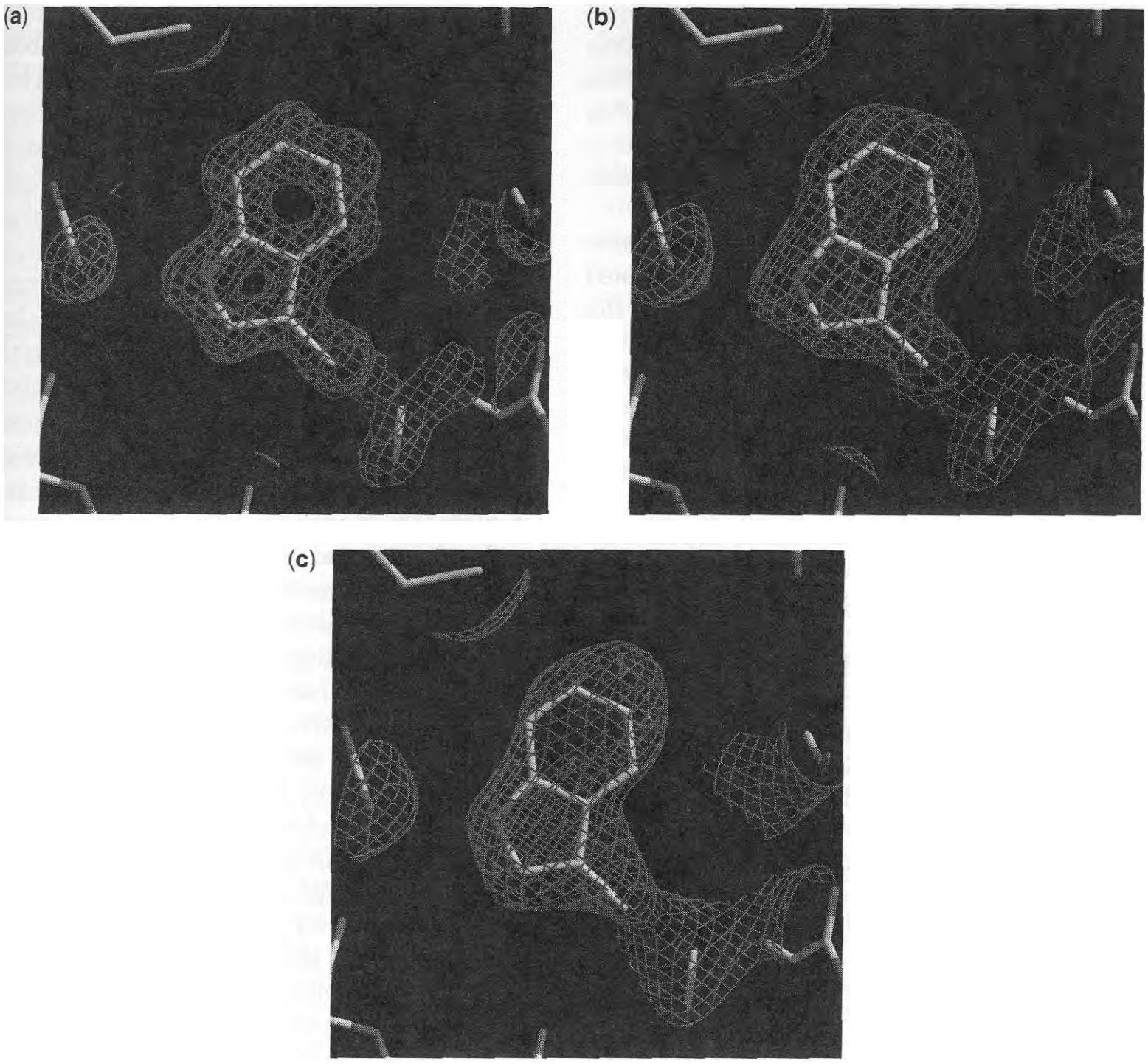


Figure 11.6. Three density maps at differing resolutions: a, **1.3** Å; b, 2.1 Å; c, 3.0 Å. See color insert.

with aver 15,000 of these done by X-ray crystallography. Most of the rest were done by NMR. For small molecules, the Cambridge Structural Database (CSD) (29) contains structural information for over 230,000 organic and organometallic compounds. All of these structures have been determined by X-ray or neutron diffraction techniques.

3 APPLICATIONS OF THE USE OF CRYSTALLOCRAPHIC STUDIES IN DRUG DISCOVERY AND DEVELOPMENT

Crystallization of small molecule compounds with a protein or nucleic acid target followed by X-ray crystallographic determination of the combined structure is the basis and hallmark of structure-based drug design. As structural biology moves into the post-genomic age, many companies and academic laboratories are faced with the challenge of co-crystallization of targets and inhibitors or activators on a scale never before attempted. Previously, crystal structure determination of a protein-substrate or inhibitor complex in an academic or industrial environment often yielded the structural information desired to understand the mechanism of action or in the design of a more suitable substrate or inhibitor. However modern day laboratories are now faced with the daunting challenge of crystallizing hundreds of compounds for clues in further ligand design using standard organic synthesis or combinatorial approaches.

A variety of methods have been employed to co-crystallize biological molecules with small molecules. Discovery of crystallization conditions is still an often tedious task, so newer methods for screening crystallization conditions for proteins include the use of semi-automated robots.

Two fundamental methods are available for co-crystallizations. One method is termed "soaking." This employs the addition of the small molecule directly to saturated solutions containing crystals of the biological macromolecule in hopes that the ligand of interest will soak directly into the crystal and bind to its active or binding site, so that the co-crystal structure can be determined. The other method, called co-crystallization, depends on having an ability to add the ligand to the aqueous protein solution in at least stoichiometric amounts, followed by crystallization using either the known crystallization conditions or by setting up a new screen for determining suitable crystallization conditions. Both methods have disadvantages and advantages, and it is primarily up to the investigator to decide which method, or both methods, should be employed for their experiments.

One limitation to the soaking method is that the amount actually dissolved and available to form the complex can often not be easily determined or controlled. In general, an excess of ligand, as a solid, is added to the solution with the crystals of protein with the hope that the ligand dissolves completely and will diffuse into the crystal binding site. One method that has demonstrated success involves equilibration of the crystal with slightly higher concentration of the crystal mother liquor that contains the ligand solubilized by organic cosolvents (i.e., isopropanol, DMSO, ethanol, etc.) as part of the medium for diffusion. However, higher levels of organic solvent often decreases the resolution of diffraction. Lowering the level of solvent after the addition of compound has been found to result in better diffracting crystals. Another major limitation of this method is that it is necessary to collect an entire X-ray diffraction data set to determine if the small molecule is bound to the protein. This trial and error method can be time-consuming or expensive if a high-throughput crystallography approach is the objective.

Co-crystallization permits highly parallel screening for bound ligands through robotic systems. The co-crystallization method is better suited for high-throughput crystallography, because ligand binding can sometimes be determined without the need to solve each structure. Faster spectral analyses, or alternatives such as native gel shifts, gel filtration, and mass spectrometry can provide information on which of the crystals should be taken into X-ray studies. One difficulty in using the co-crystallization method is the problem of determining the concentration of the protein that is most suitable for complexation. For example, if the protein solution is roughly 1 mM in high salt or aqueous buffers, many organic molecules are not as soluble at that level. In these cases a lower concentration of protein is usually employed to attain stoichiometric ratios. As described above, small percentages of organic solvent can be useful for increasing the concentration of the organic compound in solution, but not without affecting the protein stability or crystal quality. In general, lowering the protein concentration sufficiently, followed by addition of the appropriate amount of ligand, and then concentration of the mixture to the desired protein concentration for crystallization is the most successful method.

Once conditions for obtaining the complexed protein have been obtained, the next step is to decide on which crystallization conditions to use. In some cases, those proteins that do not undergo large tertiary structure changes when complexed to ligands can be crystallized under similar conditions as for the uncomplexed ligand. However, in some instances, proteins will change conformation, depending on the type of ligand that they are complexed with, and a large screening of possible new crystallization conditions is required.

In many cases, soaking a compound into a crystal is not possible because of low solubility of the compound in the aqueous mother liquor. Soaking experiments can also be limited when the conformational space of the binding site is hindered, occupied by adjacent molecules in the crystal lattice, or if there are conformational changes in the binding site because of crystal packing effects. On the other hand, co-crystallization of the protein and li-

gand, by the nature of the process, usually requires more resources in terms of protein and experimental time, leading to greater expense.

Soaking crystals and/or growing crystals in the presence of inhibitors or ligands provides an opportunity to directly observe their binding interactions, along with the often subtle conformational effects that can have a profound effect on the mode of binding. When it is possible to use this in an iterative fashion to guide the design of the next set of compounds to be synthesized in a lead series, this becomes a potent tool. Understanding how and why a compound or series binds to an active site, particularly when the affinity is also known, provides the best understanding, and the highest level at which it is possible to enable drug design. As structural biology becomes a more integrated component of drug discovery, better methods of obtaining crystal structures with and without bound ligands will be developed, with lower costs and faster turnaround times. Many companies and academic laboratories are now focused on solving these challenges. But it is also impressive to see how well we have progressed—Table 11.1 enumerates the structures of known therapeutic targets (with or without ligands) that are available in the public domain. This table is based on the Nature Biotechnology "The Usual Suspects" poster, published in 2001, but is almost identical in content to the 1997 version (30). Reference sequences for 300 of the targets were extracted from NCBI Genbank and the nonredundant database was searched with several iterations of PSI-BLAST. The resulting profile was used to search the PDB + SGX database of known structures. The top hits for each drug target are tabulated below. A great many more reside within pharmaceutical and biotech companies as proprietary structures.

4 STRUCTURAL GENOMICS

4.1 Introduction to Structural Cenomics

Until recently, structure determination by protein crystallography was a time-consuming method accessible to a few privileged skilled practitioners. X-ray crystallography was reserved to tackle questions requiring atomic resolution details of a demonstrably important protein, often a drug target. Indeed, to this day, crystallography is almost exclusively used in the pharmaceutical industry to study small molecule interactions with drug targets (see Section 3).

The development of several new methods (described in Section 2) and the availability of the complete genome sequences of both pathogens and hosts provides an unparalleled opportunity to exploit protein structures for drug discovery research in new ways. We can now contemplate using protein structure determination to help annotate genomes, that is, to assess new drug targets as well as provide multiple high-resolution structures that address selectivity issues. This emerging science of high-throughput structural biology has been termed structural genomics.

4.2 Genome Annotation

It is in infectious disease that whole genome information first became available, and it is in this field that structural genomics is having an initial impact (373). A typical approach has been to assess the viability and/or virulence of pathogens by systematic disruption of every predicted gene product. As a consequence, a large number of potential new targets have emerged: genes that are essential for pathogenicity of bacteria in a model system. Often these new genes have been filtered for those that are conserved in a variety of pathogens and that do not have a close human homolog (374). About 30–50% of the genes of a typical pathogen have no reliable functional assignment. A similar fraction of the novel targets shown to be essential also fall into this category, which becomes problematic for assay configuration. Indeed, in target-based approaches, the number of leads emerging has been disappointing. Protein structure can provide the information required to prioritize among these essential genes and to establish assays. Co-complex structures with even low-affinity hits can be used to provide key information for medicinal chemistry.

There are several ways in which structural genomics has promise as a tool for genome annotation and target prioritization. For genes of unknown function. structure can often provide clues to biochemical function. Sequence homology has become a routine method for functional assignment, but even the most

Table 11.1 Known Drug Targets with Published Structures

Target and PDB Reference	Resolution	Source	Homology	Year	Reference
Acetylcholinesterase					
1MAH(A)	3.20 Å	Green mamba	88%	1995	(31)
1B41(A), 1F8U(A)	2.76 Å, 2.90 Å	Green mamba	99%	2000	(32)
1C2B(A), 1C2O(A)	4.50 Å, 4.20 Å	Electric eel	88%	1999	(33)
1MAA(A)	2.90 Å	Mouse	88%	1998	(34)
Adenosine deaminase					
1FKX, 1FKW	2.40 Å	Mouse	82%	1996	(35)
1A4L(A), 1A4M(A)	2.60 Å, 1.95 Å	Mouse	83%	1998	(36)
1UIO, 1UIP	2.40 Å	Mouse	82%	1996	(37)
1ADD	2.40 Å	—	83%	1992	(38)
2ADA	2.40 Å	—	83%	1994	(39)
Alpha-amylase					
1JXJ(A), 1JXK(A)	1.90 Å	Human	99%	2001	(40)
1SMD	1.60 Å	Human	99%	1996	(41)
1C8Q(A)	2.30 Å	Human	99%	2000	(42)
1CPU(A), 2CPU(A)	2.00 Å	Human	97%	1999	(43)
1BSI	2.00 Å	Human	97%	1998	(44)
1HNY	1.80 Å	Human	97%	1995	(45)
3CPU(A)	2.00 Å	Human	96%	1999	(46)
1B2Y(A)	3.20 Å	Human	97%	1998	(47)
1DHK(A)	1.85 Å	Kidney bean	86%	1996	(48)
1JFH	2.03 Å	Pig	86%	1997	(49)
1PIF, 1PIG	2.30 Å, 2.20 Å	Pig	85%	1996	(50)
1OSE	2.30 Å	Pig	86%	1996	(51)
1HX0(A)	1.38 Å	Pig	86%	2001	(52)
1BVN(P)	2.50 Å	S. *tendae*	85%	1998	(53)
1PPI	2.20 Å	—	86%	1994	(54)
Androgen receptor					
1E3G(A)	2.40 Å	Human	100%	2000	(55)
1I37(A), 1I38(A)	2.00 Å	Rat	99%	2001	(56)
Anticoagulant protein C					
1AUT(C)	2.80 Å	Human	100%	1996	(57)
Aquaporin 1					
1IH5(A)	3.70 Å	Human	100%	2001	(58)
1FQY(A)	3.80 Å	Human	100%	2000	(59)

β-Amyloid					
1MWP(A)	1.80 Å	Human	100%	1999	(60)
β-Lactamase[Sa]					
1BTL	1.80 Å	Bacteria	100%	1993	(61)
1FQG(A)	1.70 Å	Bacteria	98%	2000	(62)
1JTD(A)	2.30 Å	Bacteria	99%	2001	(63)
1HTZ(A)	2.40 Å	Bacteria	98%	2001	(64)
1ERM(A), 1ERO(A)	1.70 Å, 2.10 Å	Bacteria	99%	2000	(65)
1ERQ(A)	1.90 Å	Bacteria	99%	2000	(65)
1XPB	1.90 Å	Bacteria	99%	1997	(66)
1ESU(A)	2.00 Å	Bacteria	98%	2000	(67)
1BT5(A)	1.80 Å	*Escherichia coli*	100%	1998	(68)
1TEM	1.95 Å	*Escherichia coli*	100%	1996	(69)
1CK3(A)	2.28 Å	*Escherichia coli*	99%	1999	(70)
1AXB	2.00 Å	*Escherichia coli*	100%	1997	(71)
β-Tubulin					
1JFF(B)	3.50 Å	Bovine	99%	2001	(72)
1TUB(B)	3.70 Å	Pig	99%	1997	(73)
1FFX(B)	3.95 Å	Rat	99%	2000	(74)
Calcineurin A					
1TCO(A)	2.50 Å	Bovine	99%	1996	(75)
1AUI(A)	2.10 Å	Human	99%	1997	(76)
Carbonic anhydrase 2					
1HEA, 4CAC, 5CAC	2.00 Å, 2.20 Å	Human, HSV-1	100%	1992	(30)
1G6V(A)	3.50 Å	Arabian camel	100%	2000	(77)
1CNW, 1CNX, 1CNY	2.00 Å, 1.90 Å, 2.30 Å	Human	100%	1995	(78)
1IF4(A), 1IF5(A), 1IF6(A)	1.93 Å, 2.00 Å, 2.09 Å	Human	100%	2001	(79)
1IF9(A)	2.00 Å	Human	100%	2001	(80)
1CA3, 1HEB, 1HED	2.30 Å, 2.00 Å	Human	100%	1992	(81)
1DCA, 1DCB	2.20 Å, 2.10 Å	Human	99%	1993	(82)
1CRA	1.90 Å	Human	100%	1992	(83)
1CIL, 1CIM, 1CIN	1.60 Å, 2.10 Å	Human, HSV-1	100%	1993	(84)
1CAY	2.10 Å	Human	100%	1993	(85)
1RZA, 1RZB, 1RZC, 1RZD, 1RZE	1.80 Å–1.90 Å	Human	100%	1993	(86)
2CA2	1.90 Å	Human	100%	1989	(87)
1BN1, 1BN3, 1BN4, 1BNM	2.10 Å, 2.20 Å, 2.60 Å	Human	100%	1998	(88)
1CAH	1.88 Å	Human	100%	1992	(89)

Table *11.1 (Continued)*

Target and PDB Reference	Resolution	Source	Homology	Year	Reference
1I8Z(A)	1.93 Å	Human	100%	2001	(90)
1BV3(A)	1.85 Å	Human	100%	1998	(91)
12CA	2.40 Å	Human	99%	1991	(92)
1G53(A)	1.94 Å	Human	100%	2000	(93)
1AM6	2.10 Å	Human	100%	1997	(94)
1CAN, 1CAO	1.90 Å	Human rhinovirus	100%	1992	(95)
1G0E(A), 1G0F(A)	1.60 Å	Human	99%	2000	(96)
1AVN	2.00 Å	Human	100%	1997	(97)
1UGF	2.00 Å	Human	99%	1996	(98)
1HVA	2.30 Å	Human	99%	1992	(99)
5CA2	2.10 Å	Human	99%	1991	(100)
1HCA	2.30 Å	—	100%	1992	(101)
4CA2, 6CA2, 7CA2, 9CA2	2.10 Å–2.80 Å	Human	100%	1991	(102)
1ZNC(A)	2.80 Å	Human	100%	1996	(103)
Catechol methyltransferase					
1VID	2.00 Å	Rat	80%	1996	(104)
Cholecystokinina receptor					
1D6G(A)	NMR	—	95%	1999	(105)
Coagulation factor 10					
1EZQ(A), 1F0R(A), 1F0S(A)	2.20 Å, 2.10 Å	Human	100%	2000	(106)
1C5M(D)	1.95 Å	Human	99%	1999	(107)
1XKA(C), 1XKB(C)	2.30 Å, 2.40 Å	Human	100%	1998	(108)
1FAX(A)	3.00 Å	Human	98%	1996	(109)
1FJS(A)	1.92 Å	Human	100%	2000	(110)
1KIG(H)	3.00 Å	Soft tick	83%	1997	(111)
1HCG(A)	2.20 Å	—	100%	1993	(112)
Coagulation factor 2					
1AI8(H)	1.85 Å	*Hirudo rnedicinalis*	100%	1997	(30)
1MKW(K), 1MKX(K)	2.30 Å, 2.20 Å	Bos taurus	84%	1997	(113)
1BTH(H)	2.30 Å	Bovine	99%	1996	(114)
1HXF(H)	2.10 Å	*Hirudo medicinalis*	100%	1996	(115)
1G30(B)	2.00 Å	*Hirudo medicinalis*	100%	2000	(116)
1A3E(H)	1.85 Å	*Hirudo medicinalis*	100%	1998	(117)
1D3P(B), 1D3Q(B)	2.10 Å, 2.90 Å	*Hirudo rnedicinalis*	100%	1999	(118)
1HDT(H)	2.60 Å	*Hirudo medicinalis*	100%	1994	(119)
1AD8(H)	2.00 Å	*Hirudo medicinalis*	100%	1997	

1LHC(H), 1LHF(H), 1LHG(H)	1.95 Å, 2.40 Å, 2.25 Å	***Hirudomedicinalis***	100%	1994	(121)
1JOU(B)	1.80 Å	Human	99%	2001	(122)
1DIT(H)	2.30 Å	Human	100%	1995	(123)
1UVS(H)	2.80 Å	Human	100%	1996	(124)
4THN(H)	2.50 Å	Human	100%	1998	(125)
1THP(B)	2.10 Å	Human	99%	1999	(126)
1JOU(B)	1.80 Å	Human	99%	2001	(122)
1DIT(H)	2.30 Å	Human	100%	1995	(123)
1UVS(H)	2.80 Å	Human	100%	1996	(124)
4THN(H)	2.50 Å	Human	100%	1998	(125)
1THP(B)	2.10 Å	Human	99%	1999	(126)
1AY6(H)	1.80 Å	Human	100%	1997	(127)
1C1U(H), 1C1V(H)	1.75 Å, 1.98 Å	Human	100%	1999	(128)
1A4W(H)	1.80 Å	Human	100%	1998	(129)
1G37(A)	2.00 Å	Human	100%	2000	(130)
1EOJ(A), 1EOL(A)	2.10 Å	Human	100%	2000	(131)
1BB0(B)	2.10 Å	Human	100%	1998	(132)
1C4U(2), 1D6W(A), 1D9I(A), 1DOJ(A)	1.70 Å–2.30 Å	Human	100%	1999	(133)
7KME(H)	2.10 Å	Human	100%	1999	(134)
1QBV(H)	1.80 Å	Human	100%	1999	(135)
1DM4(B)	2.50 Å	Human	99%	1999	(136)
1UMA(H)	2.00 Å	Medicinal leech	100%	1996	(137)
1BMM(H), 1BMN(H)	2.60 Å, 2.80 Å	Medicinal leech	100%	1995	(138)
1A2C(H)	2.10 Å	M. *aeruginosa*	100%	1997	(139)
1FPC(H)	2.30 Å	—	100%	1994	(140)
1NRO(H), 1NRR(H)	3.10 Å, 2.40 Å	—	100%	1994	(141)
1HAG(E)	2.00 Å	—	100%	1994	(142)
1HLT(H)	3.00 Å	—	100%	1994	(143)
1TMU(H)	2.50 Å	—	100%	1994	(144)
4HTC(H)	2.30 Å	—	100%	1993	(145)
1AIX(H), 1DWB(H), 1DWC(H)	2.10 Å, 3.16 Å, 3.00 Å	***Hirudo medicinalis***	100%	1992	(146)
2HPP(H)	3.30 Å	—	100%	1993	(147)
1ABI(H)	2.30 Å	—	100%	1992	(148)
Coagulation factor 7					
1JBU(H)	2.00 Å	Bacteria	100%	2001	(149)
Coagulation factor 7a					
1QFK(H)	2.80 Å	Human	100%	1999	(150)
1DVA(H)	3.00 Å	Human	100%	2000	(151)

Table 11.1 (Continued)

Target and PDB Reference	Resolution	Source	Homology	Year	Reference
1DAN(H)	2.00 Å	Human	100%	1997	(152)
1CVW(H)	2.28 Å	Human	100%	1999	(153)
1FAK(H)	2.10 Å	Human	100%	1998	(154)
Coagulation factor 9					
1RFN(A)	2.80 Å	Human	100%	1999	(155)
1PFX(C)	3.00 Å	Pig	88%	1995	(156)
Cox-1					
1DIY(A)	3.00 Å	Sheep	93%	1999	(157)
1CQE(A), 1PRH(A)	3.10 Å, 3.50 Å	Sheep	92%	1994	(158)
1PTH	3.40 Å	Sheep	92%	1995	(159)
1EBV(A)	3.20 Å	Sheep	93%	2000	(160)
1FE2(A)	3.00 Å	Sheep	92%	2000	(161)
1EQG(A), 1EQH(A), 1HT5(A), 1HT\ill\(A)	2.61 Å–2.75 Å	Sheep	92%	2000	(162)
1PGE(A), 1PGF(A), 1PGG(A)	3.50 Å, 4.50 Å	Sheep	92%	1995	(163)
Cox-2					
1CVU(A), 1DDX(A)	2.40 Å, 3.00 Å	Mouse	87%	1999	(164)
1CX2, 3PGH, 4COX, 5COX, 6COX	3.00 Å	Mouse	87%	1996	(165)
Cytochrome P450 **reductase**					
1B1C(A)	1.93 Å	Human	100%	1998	(166)
1AMO(A)	2.60 Å	Rat	93%	1997	(167)
1J9Z(A), 1JA0(A), 1JA1(A)	2.70 Å, 2.60 Å, 1.80 Å	Rat	92%	2001	(168)
Dihydrofolate reductase					
1BOZ(A)	2.10 Å	Human	99%	1998	(169)
1HFP, 1HFQ, 1HFR	2.10 Å	Human	100%	1997	(170)
1OHJ, 1OHK	2.50 Å	Human	100%	1997	(171)
1DR1, 1DR5, 1DR6, 1DR7	2.20 Å, 2.40 Å	—	75%	1992	(172)
1DR2, 1DR3	2.30 Å	—	75%	1992	(173)
1DR4	2.40 Å	—	75%	1992	(174)
1DHF(A), 2DHF(A)	2.30 Å	—	100%	1989	(175)
1DLR, 1DLS	2.30 Å	—	99%	1995	(176)
8DFR	1.70 Å	—	75%	1989	(177)
1DRF	2.00 Å	—	100%	1990	(178)
Dihydroorotate dehydrogenase					
1D3G(A), 1D3H(A)	1.60 Å, 1.80 Å	Human	100%	1999	(179)

Dihydropteroate synthetase[Sa]					
1AD1(A), 1AD4(A)	2.20 Å, 2.40 Å	*S. aureus*	100%	1997	(180)
DNA helicase pcra[Sa]					
1QHH(A)	2.50 Å	*B. thermophilus*	71%	1999	(181)
DNA topoisomerase 1					
1EJ9(A)	2.60 Å	Human	99%	2000	(182)
1A36(A)	2.80 Å	Human	99%	1998	(183)
1A31(A), 1A35(A)	2.10 Å, 2.50 Å	Human	90%	1998	(184)
Estrogen receptor l a					
1QKT(A), 1QKU(A)	2.20 Å, 3.20 Å	Human	98%	1999	(185)
1HCP	NMR	Human	100%	1993	(186)
1A52(A)	2.80 Å	Human	99%	1998	(187)
1ERR(A), 1ERE(A)	2.60 Å, 3.10 Å	Human	99%	1997	(188)
1HCQ(A)	2.40 Å	Human	100%	1993	(189)
3ERT(A), 3ERD(A)	1.90 Å, 2.03 Å	Human	98%	1999	(190)
FK506-binding protein					
1TCO(C)	2.50 Å	Bovine	99%	1996	(75)
1FKD, 2FKE	1.72 Å	Human	100%	1993	(191)
1FKJ, 1FKK, 1FKL	1.70 Å, 2.20 Å	Cow	97%	1995	(193)
1FAP(A)	2.70 Å	Human	100%	1996	(194)
3FAP(A), 4FAP(A)	1.85 Å, 2.80 Å	Human	100%	1999	(195)
1NSG(A)	2.20 Å	Human	100%	1997	(196)
1FKR, 1FKS, 1FKT	NMR	Human	100%	1992	(197)
1EYM(A)	2.00 Å	Human	99%	2000	(198)
1BL4(A)	1.90 Å	Human	99%	1998	(199)
1D6O(A), 1D7H(A), 1D7I(A), 1D7J(A)	1.85 Å–1.90 Å	Human	100%	1999	(200)
1QPF(A), 1QPL(A)	2.50 Å, 2.90 Å	Human	100%	1999	(201)
1F40(A)	NMR	Human	100%	2000	(202)
1B6C(A)	2.60 Å	Human	100%	1999	(203)
1A7X(A)	2.00 Å	Human	100%	1998	(204)
1BKF	1.60 Å	Human	98%	1995	(205)
2FAP(A)	2.20 Å	Human	100%	1998	(206)
1C9H(A)	2.00 Å	Human	83%	1999	(207)
1FKG, 1FKH, 1FKI(A)	2.00 Å, 1.95 Å, 2.20 Å	—	100%	1993	(208)
1FKF	1.70 Å	—	100%	1991	(209)
1FKB	1.70 Å	—	100%	1992	(210)

Table 11.1 (Continued)

Target and PDB Reference	Resolution	Source	Homology	Year	Reference
Follicle stimulating hormone					
1FL7(B)	3.00 Å	Human	99%	2000	(211)
GABA transferase					
1GTX(A)	3.00 Å	Pig	94%	1999	(212)
Glucocorticoid receptor					
1LAT(A)	1.90 Å	Rat	85%	1995	(213)
1GLU(A)	2.90 Å	—	94%	1992	(214)
Glutamate receptor 1					
1EWK(A), 1EWT(A), 1EWV(A)	2.20 Å, 3.70 Å, 4.00 Å	Rat	98%	2000	(215)
Glutathione peroxidase					
1GP1(A)	2.00 Å	—	90%	Jun 1985	(216)
G-CSF 3					
1CD9(A), 1PGR(A)	2.80 Å, 3.50 Å	Mouse	98%	1999	(217)
1BGC, 1BGD, 1BGE(A)	1.70 Å, 2.30 $2.20 Å	—	80%	1993	(218)
1GNC	NMR	—	100%	1994	(219)
1RHG(A)	2.20 Å	—	98%	1993	(220)
Granulocyte-macrophage CSF					
1CSG(A)	2.70 Å	—	100%	1992	(30)
2GMF(A)	2.40 Å	Human	100%	1996	(221)
Growth hormone receptor					
1A22(B)	2.60 Å	Human	100%	1998	(222)
1AXI(B)	2.10 Å	Human	98%	1997	(223)
1HWG(B), 1HWH(B)	2.50 Å, 2.90 Å	Human	100%	1996	(224)
3HHR(B)	2.80 Å	—	100%	1993	(225)
HIV reverse transcriptase					
1DLO	2.70 Å	HIV-1	98%	1996	(226)
1RT3(B)	3.00 Å	HIV-1	99%	1998	(227)
1HPZ, 1HQE, 1HQU	3.00 Å, 2.70 Å	HIV-1	98%	2000	(228)
1BQM, 1BQN	3.10 Å, 3.30 Å	HIV-1	98%	1998	(229)
1TVR(B), 1UWB(B)	3.00 Å, 3.20 Å	HIV-1	99%	1996	(230)
1EET	2.73 Å	HIV-1	98%	2000	(231)
1IKV, 1IKW, 1IKX, 1IKY	2.80 Å – 3.00 Å	HIV-1	98%	2001	(232)
1HVU(B)	4.75 Å	HIV-1	99%	1998	(233)

1C1B(B)	2.50 Å	HIV-1 pol	99%	1999	(234)
2HMI(B)	2.80 Å	virus	99%	1998	(235)
1HYS	3.00 Å	Virus	98%	2001	(236)
1HMV	3.20 Å	Virus	98%	1994	(237)
1HNI	2.80 Å	Virus	98%	1995	(238)
1HNV	3.00 Å	virus	98%	1995	(239)
1FKP(B)	2.90 Å	virus	99%	2000	(240)
1JLA, 1JLB, 1JLC, 1JLE, 1JLF, 1JLG	2.50 Å – 3.00 Å	virus	99%	2001	(241)
1J5O, 1QE1(B)	3.50 Å, 2.85 Å	HIV-1	99%	1999	(242)
3HVT(B)	2.90 Å	—	99%	1994	(243)
Inosine monophosphate dehydrogenase 2					
1JR1(A)	2.60 Å	Chinese hamster	98%	2001	(244)
1B3O(A)	2.90 Å	Human	100%	1998	(245)
Insulin-like growth factor 1					
3LRI(A)	NMR	Human	86%	1999	(246)
1BQT	NMR	Human	100%	1998	(247)
1IMX(A)	1.82 Å	Human	100%	2001	(248)
1B9G(A)	NMR	—	81%	1999	(249)
2GF1, 3GF1	NMR	—	100%	1991	(250)
Insulin-like growth factor 1 receptor					
1IGR(A)	2.60 Å	Human	100%	1998	(251)
1GAG(A)	2.70 Å	Human	78%	2000	(252)
1I44(A)	2.40 Å	Human	79%	2001	(253)
1IR3(A)	1.90 Å	Human	78%	1997	(254)
1IRK	2.10 Å	—	79%	1995	(255)
Insulin-like growth factor 2					
1IGL	NMR	—	100%	1994	(256)
Integrin alpham					
1BHQ(1), 1IDN(1)	2.70 Å	Human	100%	1998	(257)
1JLM	2.00 Å	Human	100%	1996	(258)
1IDO	1.70 Å	Human	96%	1996	(259)
Intercellular adhesion molecule 1					
1IAM	2.10 Å	Human	98%	1998	(260)
1IC1(A)	3.00\?\	Human	100%	1998	(261)
1D3E(I), 1D3I(I), 1D3L(A)	2.80 Å, 2.60 Å, 3.25 Å	Human rhinovirus	100%	1999	(262)

Table 11.1 (Continued)

Target and PDB Reference	Resolution	Source	Homology	Year	Reference
Interferon a 1					
1ITF	NMR	Human	82%	1997	(263)
1RH2(A)	2.90 Å	Human	83%	1996	(264)
Interferon γ					
1FG9(A)	2.90 Å	Human	100%	2000	(265)
1FYH(A)	2.04 Å	Human	100%	2000	(266)
1EKU(A)	2.90 Å	Human	98%	2000	(267)
1HIG(A)	3.50 Å	—	99%	1991	(268)
Interleukin 1					
2ILA	2.30 Å	—	99%	1991	(269)
Interleukin 1 receptor					
1G0Y(R)	3.00 Å	Human	100%	2000	(270)
1IPA(Y)	2.70 Å	Human	100%	1998	(271)
1ITB(B)	2.50 Å	Human	100%	1997	(272)
Interleukin 10					
1VLK	1.90 Å	Epstein-Barr virus	92%	1997	(273)
21LK	1.60 Å	Human	100%	1996	(274)
1ILK	1.80 Å	Human	100%	1995	(275)
1J7V(L)	2.90 Å	Human	100%	2001	(276)
1INR	2.00 Å	Human	100%	1995	(277)
Interleukin 12					
1F42(A), 1F45(A)	2.50 Å, 2.80 Å	Human	100%	2000	(278)
Interleukin 13					
1GA3(A)	NMR	Human	100%	2000	(279)
Interleukin 2					
1IRL	NMR	Human	99%	1995	(280)
3INK(C)	2.50 Å	—	99%	1992	(281)
Interleukin 3					
1JLI	NMR	*Escherichia coli*	87%	1995	(282)
Interleukin 4					
1HIJ, 1HIK	3.00 Å, 2.60 Å	Human	100%	1995	(283)
1IAR(A)	2.30 Å	Human	100%	1999	(284)
1HZI(A)	2.05 Å	Human	99%	2001	(285)
1ITM	NMR	—	100%	1994	(286)
1BBN, 1BCN	NMR	—	98%	1992	(287)

1CYL	NMR	—	100%	1994	(288)
2CYK	NMR	—	100%	1994	(289)
1ITL	NMR	—	100%	1992	(290)
21NT	2.40 Å	—	100%	1993	(291)
1RCB	2.25 Å	—	100%	1992	(292)
1ITI	NMR	—	98%	1993	(293)
Interleukin 5					
1HUL(A)	2.40 Å	Human	100%	1995	(294)
Interleukin 6					
1IL6, 2IL6	NMR	Human	100%	1997	(295)
1ALU	1.90 Å	Human	100%	1997	(296)
Interleukin 8					
1IKL, 1IKM	NMR	Human	100%	1995	(297)
1ICW(A)	2.01 Å	Human	97%	1996	(298)
1ILP(A), 1ILQ(A)	NMR	Human	100%	1998	(299)
1QE6(A)	2.35 Å	Human	97%	1999	(300)
1ROD(A)	NMR	Human	79%	1995	(301)
3IL8	2.00 Å	—	100%	1990	(302)
1IL8(A), 2IL8(A)	NMR	—	100%	1990	(303)
Leukotriene A4 hydrolase					
1HS6(A)	1.95 Å	Human	100%	2000	(304)
Lipocortin I					
1AIN	2.50 Å	Human	100%	1992	(305)
1HM6(A)	1.80 Å	Pig	89%	2000	(306)
Luteinizing hormone β					
1QFW(B)	3.50 Å	*Escherichia coli*	83%	1999	(307)
1HCN(B)	2.60 Å	—	83%	1994	(308)
1HRP(B)	3.00 Å	—	83%	1994	(309)
Macrophage CSF 1					
1HMC(A)	2.50 Å	—	100%	1993	(310)
Neuraminidase[int B virus]					
1INF	2.40 Å	Influenza b virus	100%	1995	(311)
1A4G(A), 1A4Q(A)	2.20 Å, 1.90 Å	Influenza b virus	94%	1998	(312)
1IVB	2.40 Å	—	99%	1994	(313)
1NSC(A), 1NSD(A)	1.70 Å, 1.80 Å	—	94%	1993	(314)
1B9S(A), 1B9T(A), 1B9V(A)	2.50 Å, 2.40 Å, 2.35 Å	Influenza b virus	99%	1999	(315)
1INV	2.40 Å	—	99%	1994	(316)
1NSB(A)	2.20 Å	—	94%	1991	(317)

Table 11.1 (Continued)

Target and PDB Reference	Resolution	Source	Homology	Year	Reference
Neuropeptide Y					
1RON	NMR	Human	100%	1996	(318)
1F8P(A)	NMR	—	97%	2000	(319)
1FVN(A)	NMR	—	91%	2000	(320)
Parathyroid hormone					
1HTH	NMR	Human	88%	1997	(321)
1FVY(A)	NMR	Human	100%	2000	(322)
1BWX, 1HPY, 1ZWA, 1ZWC	NMR	Human	100%	1998	(323)
1ET1(A)	0.90 Å	Human	100%	2000	(324)
1HPH	NMR	—	100%	1995	(325)
1ZWB, 1ZWD, 1ZWE, 1ZWF, 1ZWG	NMR	—	100%	1996	(326)
PDGF β					
1PDG(A)	3.00 Å	—	100%	1992	
Phospholipase A2					
1BCI	NMR	Human	100%	1998	(327)
1RLW	2.40 Å	Human	98%	1997	(328)
1CJY(A)	2.50 Å	Human	100%	1999	(329)
Potassium channel shaker					
1A68	1.80 Å	Sea hare	87%	1998	(330)
1EOD(A), 1EOE(A), 1EOF(A)	2.45 Å, 1.70 Å, 2.38 Å	Sea hare	87%	2000	(331)
1T1D(A)	1.51 Å	Sea hare	88%	1998	(332)
1EXB(E)	2.10 Å	Rat	100%	2000	(333)
1DSX(A), 1QDV(A), 1QDW(A)	1.60 Å, 2.10 Å	Rat	94%	2000	(334)
PPAR γ					
4PRG(A)	2.90 Å	Escherichia coli	97%	1999	(335)
1PRG(A), 2PRG(A)	2.20 Å, 2.30 Å	Human	97%	1998	(336)
1FM6(D), 1FM9(D)	2.10 Å	Human	99%	2000	(337)
3PRG(A)	2.90 Å	Human	99%	1998	(338)
Progesterone receptor					
1E3K(A)	2.80 Å	Human	100%	2000	(55)
1A28(A)	1.80 Å	Human	100%	1998	(339)
Prolactin receptor					
1BP3(B)	2.90 Å	Human	100%	1998	(340)
1F6F(B)	2.30 Å	Rat	72%	2000	(341)
Retinoic acid receptor					
1DSZ(A)	1.70 Å	Human	100%	2000	(342)
1EXA(A), 1EXX(A)	1.59 Å, 1.67 Å	Human	81%	2000	(343)

2LBD	2.00 Å	Human	80%	1997	(344)
3LBD, 4LBD	2.40 Å	Human	80%	1998	(345)
1DKF(B)	2.50 Å	Human	100%	1999	(346)
1FCX(A), 1FCY(A), 1FCZ(A)	1.47 Å, 1.30 Å, 1.38 Å	Human	83%	2000	(347)
1HRA	NMR	—	94%	1993	(348)
Retinoid X receptor					
1FM6(A), 1FM9(A)	2.10 Å	Human	100%	2000	(337)
1DSZ(B)	1.70 Å	Human	100%	2000	(342)
1DKF(A), 1LBD	2.50 Å, 2.70 Å	Human	99%	1999	(349)
2NLL(A)	1.90 Å	Human	100%	1996	(350)
1RXR	NMR	Human	98%	1998	(351)
1G1U(A), 1G5Y(A)	2.50 Å, 2.00 Å	Human	100%	2000	(352)
1FBY(A)	2.25 Å	Human	100%	2000	(353)
1BY4(A)	2.10 Å	Human	100%	1998	(354)
Serotransferrin β					
1JNF(A)	2.60 Å	Rabbit	78%	2001	(355)
Stem cell factor					
1EXZ(A)	2.30 Å	Human	97%	2000	(356)
1SCF(A)	2.20 Å	Human	88%	1998	(357)
Thymidine kinase[HHV]					
1KIM(A)	2.14 Å	HSV-1	98%	1997	(358)
1OHI(A), 2KI5(A)	1.90 Å	HSV-1	98%	1999	(359)
1KI2, 1KI3, 1KI4, 1KI6, 1KI7, 1KI8	2.20 Å – 2.37 Å	HSV-1	99%	1998	(360)
1VTK, 2VTK, 3WK	2.75 Å, 2.80 Å, 3.00 Å	HSV-1	98%	1997	(361)
1E2H(A), 1E2I(A), 1E2J(A)	1.90 Å, 2.50 Å	—	99%	2000	(362)
1E2M(A), 1E2N(A), 1E2P(A)	2.20 Å, 2.50 Å	HSV-1	99%	2000	(363)
1E2K(A), 1E2L(A)	1.70 Å, 2.40 Å	—	99%	2000	(364)
Tumor necrosis factor receptor 1					
1NCF(A)	2.25 Å	Human	100%	1994	(365)
1EXT(A)	1.85 Å	Human	100%	1996	(366)
1TNR(R)	2.85 Å	—	100%	1994	(367)
Vitamin D receptor					
1IE8(A), 1IE9(A)	1.52 Å, 1.40 Å	Human	83%	2001	(368)
1DB1(A)	1.80 Å	Human	83%	1999	(369)
Xanthine-guanine phosphoribosyltransferase					
1A95(A), 1A97(A), 1A98(A)	2.00 Å, 2.60 Å, 2.25 Å	*Escherichia coli*	100%	1998	(370)
1NUL(A)	1.80 Å	*Escherichia coli*	100%	1996	(371)
1A96(A)	2.00 Å	*Escherichia coli*	100%	1998	(372)

sensitive sequence methods [such as ISS (375)] fail to identify many homologous relationships. Structure is more conserved than sequence, so structural classification schemes (SCOP, CATH) have been a valuable method to assign proteins to functional groups. A now classic example of functional understanding from structural homology was the discovery that the Bcl-2 family of apoptosis proteins are homologous to pore-forming toxins (376). This finding led to the suggestion that Bcl-2 proteins may function by perforating mitochondrial membranes, and has since opened new avenues of fruitful research.

In addition to structural homology as assessed by global similarities, local structural features can give clues to structure even when proteins are not homologous. By identifying surface clusters of polar residues that are well conserved in the sequence family it is possible to identify likely functional sites even when there is no obvious structural homology. These three-dimensional motifs can be compared with a structure database to identify similar motifs with known function. A classic example is found among the serine proteases. Chymotrypsin and subtilisin share a similar catalytic triad (His-Asp-Ser) but are otherwise unrelated structurally. The PLP-dependent enzymes are famed for the diversity of both structure and function, but even among this group, common structural motifs seem to have evolved convergently (377).

Simply searching for large clefts in the protein surface turns out to be an extremely successful method to identify active sites. Nucleic acid binding functions can be particularly obvious from an analysis of surface electrostatics (378,379). Mice homozygous for tubby loss-of-function mutations show an obese phenotype, and therefore, the tubby protein has attracted considerable interest. However, 3 years after the initial cloning of the tubby gene (380,3811, the molecular function of the protein was still a mystery. The structure of the conserved C-terminal domain of tubby was determined by X-ray crystallography, and a large groove of highly positive charge immediately led to the hypothesis that the protein acted as a transcription factor (6). The search is now on for downstream targets of tubby, and further structural work has demonstrated a new role for the tubby protein in G-protein–coupled receptor (GPCR)-mediated signal transduction (382) This on-going story demonstrates the power of structural approaches to determine function.

In tackling structures of proteins of unknown function bound metal ions, natural substrates or even serendipitously bound small molecules arising from the crystal preparation (e.g., buffers) often suggests the location of an active site. If the side-chains contributing to binding are well conserved, then this is good evidence locating an active site and helps assess the "drugability" of the protein. The recent structure of LuxS illustrates the power of this approach (373). A number of genes had been identified that are required for quorum sensing in bacteria by system 2. Quorum sensing by the widely conserved system 2 has emerged as an intriguing mechanism by which bacteria monitor their density and seems to be an important component of the progression to virulence, at least in certain pathogens. LuxS is the product of one of the genes required for system 2, but nothing was known of the molecular function of LuxS in this pathway. Disruption of this pathway has promise in antibacterial drug design, but whether LuxS would be an attractive target for small molecule design was unclear. No information was available to develop a biochemical assay, and besides, it was not clear what kind of library to use in high-throughput screening. The structure of LuxS was solved at Structural GenomiX in less than 2 months, and there are representative X-ray structures from three different bacteria. The structure showed that LuxS forms a dimer in which each monomer has a zinc ion coordinated by a His-His-Cys triad and water molecule. Non-covalently bound methionine molecules were found to have bound in a pocket formed at the dimer interface and close to the zinc ions (see Fig. 11.7, a and b). Methionine was shown to have bound as an artefact of the purification procedure. With this information, it became immediately clear that LuxS is likely a zinc metalloenzyme, and a hypothesis for the likely physiological substrate emerged from molecular modeling studies of the methionine binding site. This example illustrates how structure could rapidly accelerate an early stage project providing the starting point for assay development and selection of an appropriate

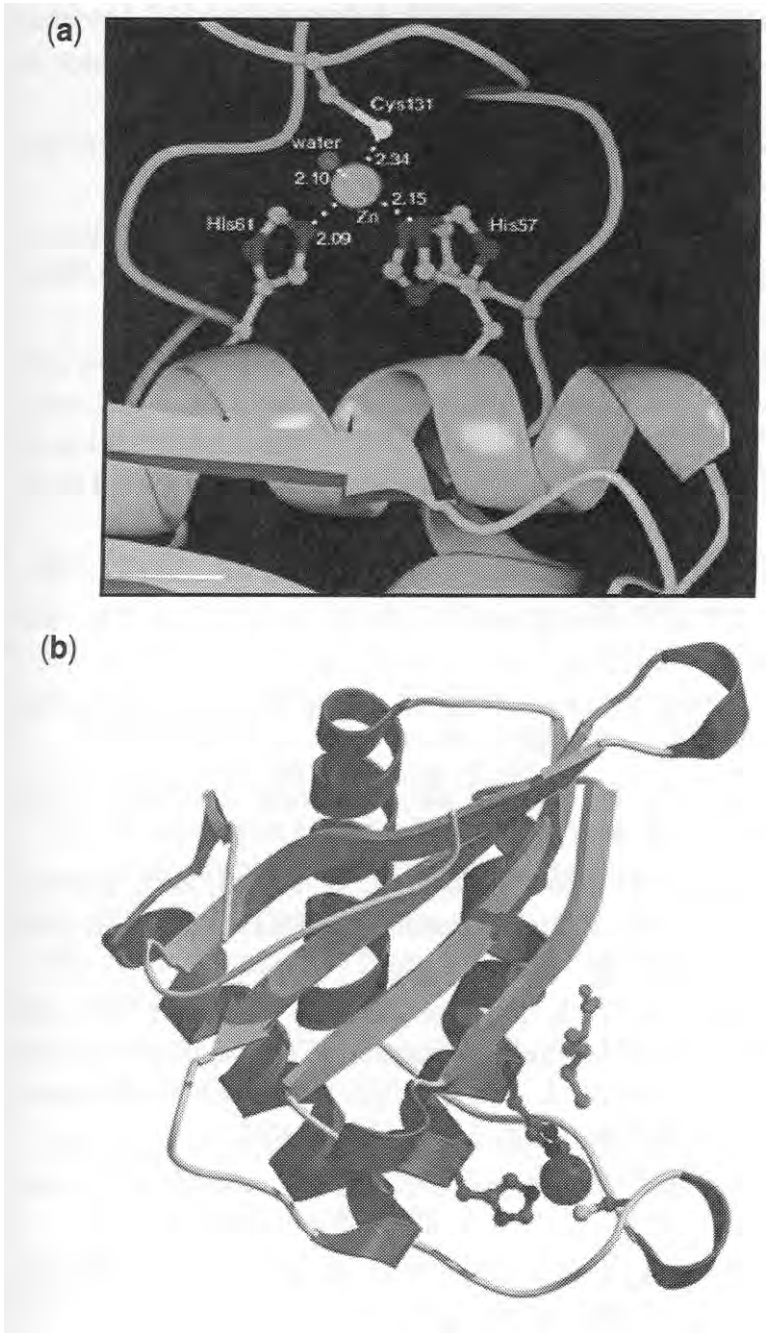


Figure **11.7.** (a) The likely active site of LuxS identified by searching for clusters of polar, conserved residues in the structure. (b) Structure of the LuxS monomer highlighting the bound zinc ion (magenta) and methionine (green). See color insert.

screening library (in this case metalloenzyme inhibitor libraries would be desirable). The model of the likely substrate bound to the active site suggests further experiments to test this hypothesis and even provides a starting point for medicinal chemistry exploration.

4.3 Pathways

Increasingly in drug discovery, particular molecular pathways are attracting interest in drug design and often manipulation of any of a number of pathway components would achieve the same end. Pathways controlling apoptosis, the cell cycle, and inflammation all contain multiple biologically validated targets. In microbial disease, several biosynthetic pathways, such as peptidoglycan biosynthesis and translation, are the targets of current drugs and several new pathways are promising targets for the development of novel agents.

Comprehensive, high-resolution structural information of multiple pathway components provides a basis for the rational design of inhibitors targeting the pathway. Interfering with the function of any of a number of enzymes of a pathway may have equally beneficial therapeutic value. Despite this, some enzymes may be more tractable targets for the design of inhibitors than others. Access to high resolution structural information of all the components of a therapeutically relevant pathway enables the rational choice of the best-suited target(s) to pursue for the design of agonists and antagonists. This choice may depend on such pragmatic considerations as the access to libraries targeted to particular enzyme types and available synthetic chemistry expertise. Furthermore, comparison of the binding pockets of consecutive enzymes in the pathway that bind similar (or identical) substrates and products may even enable the design of inhibitors of multiple pathway components. Such a compound may be particularly desirable in the development of novel anti-microbial and cancer agents where compound resistance can rapidly emerge. The evolution of resistance to a drug that inhibits two consecutive enzymes in an essential pathway is theoretically much less probable than evolution of resistance to a single enzyme inhibitor.

The non-mevalonate isopentenyl pyrophosphate biosynthesis pathway has attracted attention in recent years as a novel target for the design of anti-microbial inhibitors (383). At Structural GenomiX, the structures of three consecutive enzymes in this pathway have been solved. There is now a clear understanding of which pathway components may be most tractable to inhibitor discovery, which likely have least structural homology to human proteins, and even how to go about the design of pan-pathway inhibitors.

4.4 Protein Structure Modeling

An aim of structural genomics efforts, to provide high quality three-dimensional structures for every protein sequence, will not be achieved by experimental approaches alone

Many consortiurns are selecting targets for X-ray crystallography that would provide the templates for comparative modeling techniques of all other sequences (384, 385). As more structures are determined by NMR and X-ray crystallography, the quality of the models will improve simply because more similar templates will become available but also because and new methods for loop modeling and *ab initio* structure prediction will undoubtedly emerge (386, 387). Efforts are also underway both in industry and academia to assemble databases of homology models for all sequences that can be reasonably well modeled (388).

5 CONCLUSION

Anyone who is involved or interested in drug discovery will recognize the potential of protein crystallography to greatly enhance the process. Whether this promise has been met to date is the subject of considerable debate. What is certain, however, is that in the very near future the advances in crystallography technology will render this question moot. The histograms on the PDB website (27, 28) that show the increasing rate of structures deposited over the last decade are a startling visual indicator of the revolution that is occurring in the field. Clearly, the impact will be felt in drug discovery very soon and perhaps very dramatically, and it serves the audience of this series to be well informed of these advances in technology and their subtle limitations.

It is tempting to draw analogy with the development of other analytical technologies (NMR, FAB-MS) and conclude that protein crystallography will soon leave the incubator of "big machine physics" to become an everyday, routine tool used in the medicinal chemistry laboratory. Hopefully, this chapter has shown some of the subtle complexities of sample preparation and handling, data collection, and refinement, etc. that temper this vision and will likely keep this a specialized field for some time.

REFERENCES

1. W. A. Hendrickson, Science, 254, 51–58 (1991).
2. T. C. Tenvilliger, Nat. Struct. Biol., 7, 935-939 (2000).
3. L. Pauling, Lecture presented at the International Congress of X-ray Crystallography at Stonybrook, NY, August 1973.
4. J. Jancarik and S. H. Kim, J. *Appl. Cryst.*, 24, 409-411 (1991).
5. Hampton Research, available online at http://www.hamptonresearch.com, accessed on October 9, 2001.
6. G. L. Gilliland, M. Tung, D. M. Blakeslee, and J. Ladner, Biological Macromolecule Crystallization Database (BMCD), available online at http://www.bmal.nist.gov:8080/bmcdmal.html, accessed on October 9, 2001.
7. N. E. Chayen, Structure, 5, 1269-1274 (1997).
8. I. Jurisica, et al., *IBM Systems* J., 40, 394-409 (2001).
9. A. D'Arcy, et al., J. *Cryst.* Growth, 168, 175-180 (1996).
10. C. W. Carter Jr., Methods Enzymol., 276, 74-99 (1997).
11. A. C. Bloomer and U. W. Arndt, Acta *Crystal*logr. D Biol. *Crystallogr.*, 55(Pt 10), 1672-1680 (1999).
12. (a) G. A. Petsko, *J.* Mol. Biol., 96, 381-392 (1975); (b) R. L. Sutton, J. Chem. *Soc. Faraday* Trans., 1, 101-105 (1991); (c) D. W. Rodgers, Structure, 2, 1135-1140 (1994).
13. D. W. Green, V. M. Ingram and M. F. Perutz, Proc. R. *Soc.* A, 225, 287-307 (1954).
14. M. G. Rossman and D. M. Blow, Acta *Cryst.*, 15, 24-34 (1962).
15. L. M. Rice, T. N. Earnest, A. T. Brunger, Acta *Cryst.* D., 56, 1413-1420 (2000).
16. W. A. Hendrickson, et al., EMBO J. 9, 1665-1672 (1990).
17. W. I. Weis, et al., Science, 254, 1608-1615 (1991).
18. W. M. Clemons, Jr., et al., J Mol Biol. 310, 827-843 (2001).
19. Collaborative Computational Project, Number 4, Acta *Cryst.* D, 50, 760-763 (1994).
20. Z. Otwinowski and W. Minor, available online at http://www.hkl-xray.com, accessed October 9, 2001.
21. T. Terwilliger, Automated Structure Solution for MIR and MAD, available online at http://www.solve.lanl.gov, accessed October 9, 2001.
22. C. M. Weeks, S. A. Potter, J. Rappleye, R. Miller, available online at http://www.hwi.buffalo.edu/SnB, accessed on October 9, 2001.
23. G. M. Sheldrick, available online at http://shelx.uni-ac.gwdg.de/SHELX/, accessed on October 9, 2001.

24. V. S. Lamzin and A. Perrakis, available online at http://www.embl-hamburg.de/ARP, accessed on October 9,2001.
25. A. Jones and M. Kjeldgaard, available online at http://www.imsb.au.dk/~mok/o, accessed on October 9,2001.
26. A. T. Brunger, P. D. Adams, G. M. Clore, W. L. Delano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren, Crystallography and NMR System, available online at http://cns.csb.yale.edu/vl.0, accessed on October 9,2001.
27. H. M. Berman, D. S. Goodsell, and P. E. Bourne, Am. Scientist, 90,350–359 (2002).
28. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The Protein Data Bank, Nucleic Acids Res., 28,235–242 (2000).
29. Information on how to obtain this database is available at: http://www.ccdc.cam.ac.uW prods/.
30. J. Drews and S. Ryser, Nature *Biotechnol.* 15, (1997).
31. Y. Bourne, P. Taylor, and P. Marchot, Cell, 83, 503 (1995).
32. G. Kryger, M. Harel, M. Harel, A. Shafferman, I. Silman, and J. L. Sussman, *Acta* Crystallogr., Sect. D, 56, 1385 (2000).
33. Y. Bourne, J. Grassi, P. E. Bougis, and P. Marchot, J. Biol. Chem., 274,3370 (1999).
34. Y. Bourne, P. Taylor, P. E. Bougis, and P. Marchot, J. Biol. Chem., 274,2963 (1999).
35. V. Sideraki, K. A. Mohamedali, D. K. Wilson, Z. Chang, R. E. Kellems, F. A. Quiocho, and F. B. Rudolph, Biochemistry, 35,7862 (1996).
36. Z. Wang and F. A. Quiocho, Biochemistry, 37, 8314 (1998).
37. V. Sideraki, D. K. Wilson, L. C. Kurz, F. A. Quiocho, and F. B. Rudolph, Biochemistry, 35, 15019 (1996).
38. D. K. Wilson, F. B. Rudolph, and F. A. Quiocho, Science, 252, 1278 (1991).
39. D. K. Wilson and F. A. Quiocho, Biochemistry, 32, 1689 (1993).
40. N. Ramasubbu, C. Ragunath, and Z. Wang, In press.
41. N. Rarnasubbu, P. Venugopalan, Y. Luo, G. D. Brayer, and M. J. Levine, In press.
42. N. Ramasubbu, K. Sekar, and D. Velmurugan, *Acta* Crystallogr. D Biol. Crystallogr., 52, 435 (1996).
43. G. D. Brayer, G. Sidhu, R. Maurus, E. H. Rydberg, C. Braun, Y. Wang, N. T. Nguyen, C. M. Overall, and S. G. Withers, Biochemistry, 39, 4778 (2000).
44. E. H. Rydberg, G. Sidhu, H. C. Vo, J. Hewitt, H. C. Cote, Y. Wang, S. Numao, R. T. Macgillivray, C. M. Overall, G. D. Brayer, and S. G. Withers, Protein Sci., 8,635 (1999).
45. G. D. Brayer, Y. Luo, and S. G. Withers, Protein Sci., 4, 1730 (1995).
46. G. D. Brayer, G. Sidhu, R. Maurus, E. H. Rydberg, C. Braun, Y. Wang, N. T. Nguyen, C. M. Overall, and S. G. Withers, Biochemistry, 39, 4778 (2000).
47. M. Qian, R. Haser, G. Buisson, E. Duee, and F. Payan, Biochemistry, 33,6284 (1994).
48. C. Bompard-Gilles, P. Rousseau, P. Rouge, and F. Payan, Structure *(Lond)*, 4, 1441 (1996).
49. M. Qian, S. Spinelli, H. Driguez, and F. Payan, Protein Sci., 6, 2285 (1997).
50. M. Machius, L. Vertesy, R. Huber, and G. Wiegand, J. Mol. Biol., 260,409 (1996).
51. C. Gilles, J. P. Astier, G. Marchis-Mouren, C. Cambillau, and F. Payan, Eur. J. Biochem., 238,561 (1996).
52. M. Qian, V. Nahoum, J. Bonicel, H. Bischoff, B. Henrissat, and F. Payan, Biochemistry, 40, 7700 (2001).
53. G. Wiegand, O. Epp, and R. Huber, J. Mol. Biol., 247, 99 (1995).
54. M. Qian, R. Haser, G. Buisson, E. Duee, and F. Payan, Biochemistry, 33,6284 (1994).
55. P. M. Matias, P. Donner, R. Coelho, M. Thomaz, C. Peixoto, S. Macedo, N. Otto, S. Joschko, P. Scholz, A. Wegg, S. Basler, M. Schafer, U. Egner, and M. A. Carrondo, J. Biol. Chem., 275,26164 (2000).
56. J. S. Sack, K. F. Kish, C. Wang, R. M. Attar, S. E. Kiefer, Y. Ang. Y. Wu, J. E. Scheffler, M. E. Salvati, S. R. Krystekjr., R. Weinmann, and H. M. Einspahr, Proc. *Nat.* h a d . Sci. USA, 98, 4904 (2001).
57. T. Mather, V. Oganessyan, P. Hof, R. Huber, S. Foundling, C. Esmon, and W. Bode, Embo J., 15,6822 (1996).
58. G. Ren, V. S. Reddy, A. Cheng, P. Melnyk, and A. K. Mitra, Proc. *Nat. Acad.* Sci. USA, 98, 1398 (2001).
59. K. Murata, K. Mitsuoka, T. Hirai, T. Walz, P. Agre, J. B. Heymann, A. Engel, and Y. Fujiyoshi, Nature, 407,599 (2000).
60. J. Rossjohn, R. Cappai, S. C. Feil, A. Henry, W. J. Mckinstryd. Galatis, L. Hesse, G. Mul-

thaup, K. Beyreuther, C. L. Masters, and M. W. Parker, *Nat.* Struct. Biol., 6, 327 (1999).
61. C. Jelsch, L. Mourey, J. M. Masson, and J. P. Samama, Proteins, 16, 364 (1993).
62. N. C. Strynadka, H. Adachi, S. E. Jensen, K. Johns, A. Sielecki, C. Betzel, K. Sutoh, and M. N. James, Nature, 359, 700 (1992).
63. D. C. Lim, H. U. Park, L. Decastro, S. G. Kang, H. S. Lee, S. Jensen, K. J. Lee, and N. C. J. Strynadka, *Nat.* Struct. Biol., 8,848 (2001).
64. M. C. Orencia, J. S. Yoon, J. E. Ness, W. P. Stemmer, and R. C. Stevens, *Nat.* Struct. Biol., 8, 238 (2001).
65. S. Ness, R. Martin, A. M. Kindler, M. Paetzel, M. Gold, J. B. Jones, and N. C. J. Strynadka, Biochemistry, 39, 5312 (2000).
66. E. Fonze, P. Charlier, Y. To'Th, M. Vermeire, X. Raquet, and A. Dubus, Acta Crystallogr., Sect. D, 61,682 (1995).
67. E. Fonze, P. Charlier, Y. To'Th, M. Vermeire, X. Raquet, and A. Dubus, Acta Crystallogr., Sect. D, 51,682 (1995).
68. L. Maveyraud, L. Mourey, L. P. Kotra, J. D. Pedelacq, V. Guillet, S. Mobashery, and J. P. Samama, *J. Am.* Chem. *Soc.*, 120,9748 (1998).
69. L. Maveyraud, I. Massova, C. Birck, K. Miyashita, J. P. Samama, and S. Mobashery, J. Am. Chem. *Soc.*, 118, 7435 (1996).
70. P. Swaren, D. Golemi, S. Cabantous, A. Bulychev, L. Maveyraud, S. Mobashery, and J. P. Samama, Biochemistry, 38,9570 (1999).
71. L. Maveyraud, R. F. Pratt, and J. P. Samama, Biochemistry, 37,2622 (1998).
72. J. Lowe, H. Li, K. H. Downing, and E. Nogales, J. Mol. Biol., 313, 1045 (2001).
73. E. Nogales, S. G. Wolf, and K. H. Downing, Nature, 391, 199 (1998).
74. B. Gigant, P. A. Curmi, C. Martin-Barbey, E. Charbaut, S. Lachkar, L. Lebeau, S. Siavoshian, A. Sobel, and M. Knossow, Cell, 102, 809 (2000).
75. J. P. Griffith, J. L. Kim, E. E. Kim, M. D. Sintchak, J. A. Thomson, M. J. Fitzgibbon, M. A. Fleming, P. R. Caron, K. Hsiao, and M. A. Navia, Cell, 82, 507 (1995).
76. C. R. Kissinger, H. E. Parge, D. R. Knighton, C. T. Lewis, L. A. Pelletier, A. Tempczyk, V. J. Kalish, K. D. Tucker, R. E. Showalter, E. W. Moomaw, L. N. Gastinel, N. Habuka, X. Chen, F. Maldonado, J. E. Barker, R. Bacquet, and J. E. Villafranca, Nature, 378,641 (1995).
77. A. Desmyter, K. Decanniere, S. Muyldermans, and L. Wyns, In press.
78. P. A. Boriack, D. W. Christianson, J. Kingery-Wood, and G. M. Whitesides, J. Med. Chem., 38,2286 (1995).
79. C.-Y. Kim and D. W. Christianson, In press.
80. B. A. Grzybowski, A. V. Ishchenko, C.-Y. Kim, G. Topalov, R. Chapman, D. W. Christianson, G. M. Whitesides, and E. I. Shakhnovich, Proc. *Nat. Acad.* Sci. USA, 99,1270 (2002).
81. S. K. Nair and D. W. Christianson, Biochemistry, 32,4506 (1993).
82. J. A. Ippolito and D. W. Christianson, Biochemistry, 32,9901 (1993).
83. S. Mangani and A. Liljas, J. Mol. Biol., **232**, 9 (1993).
84. G. M. Smith, R. S. Alexander, D. W. Christianson, B. M. McKeever, G. S. Ponticello, J. P. Springer, W. C. Randall, J. J. Baldwin, and C. N. Habecker, Protein Sci., 3,118 (1994).
85. K. Hakansson, C. Briand, V. Zaitsev, Y. Xue, and A. Liljas, Acta Crystallogr. D Biol. *Crystal*logr., 50, 101 (1994).
86. K. Hakansson, A. Wehnert, and A. Liljas, Acta Crystallogr. D Biol. Crystallogr., **50**, 93 (1994).
87. A. E. Eriksson, P.M. Kylsten, T. A. Jones, and A. Liljas, Proteins, 4,283 (1988).
88. P. A. Boriack-Sjodin, S. Zeitlin, H. H. Chen, L. Crenshaw, S. Gross, A. Dantanarayana, P. Delgado, J. A. May, T. Dean, and D. W. Christianson, Protein Sci., 7, 2483 (1998).
89. K. Hakansson and A. Wehnert, J. Mol. Biol., 228, 1212 (1992).
90. C.-Y. Kim, D. A. Whittington, J. S. Chang, J. Liao, J. A. May, and D. W. Christianson, J. Med. Chem., 45, 888 (2002).
91. F. Briganti, S. Mangani, A. Scozzafava, G. Vernaglione, and C. T. Supuran, J. Biol. *Inorg.* Chem., 4, 528 (1999).
92. S. K. Nair, T. L. Calderone, D. W. Christianson, and C. A. Fierke, J. Biol. Chem., 266, 17320 (1991).
93. C.-Y. Kim, J. S. Chang, J. B. Doyon, T. T. Bairdjr., C. A. Fierke, A. Jain, and D. W. Christianson, J. Am. Chem. *Soc.*, 122,12125 (2000).
94. L. R. Scolnick, A. M. Clements, J. Liao, L. Crenshaw, M. Hellberg, J. May, T. R. Dean, and D. W. Christianson, *J. Am.* Chem. *Soc.*, 119,850 (1997).
95. S. Mangani and K. Hakansson, Eur. J. *Biochem.*, 210,867 (1992).
96. D. Duda, C. Tu, M. Qian, P. Laipis, M. Agbandje-Mckenna, D. N. Silverman, and R. Mckenna, Biochemistry, 40,1741 (2001).
97. F. Briganti, S. Mangani, P. Orioli, A. Scozzafava, G. Vernaglione, and C. T. Supuran, Biochemistry, 36,10384 (1997).
98. L. R. Scolnick and D. W. Christianson, Biochemistry, 35, 16429 (1996).

99. R. S. Alexander, L. L. Kiefer, C. A. Fierke, and D. W. Christianson, Biochemistry, **32, 1510 (1993).**

100. J. F. Krebs, C. A. Fierke, R. S. Alexander, and D. W. Christianso, Biochemistry, **30, 9153 (1991).**

101. S. K. Nair and D. W. Christianson, ***J. Am.*** Chem. *Soc.*, **113,9455 (1991).**

102. R. S. Alexander, S. K. Nair, and D. W. Christianson, Biochemistry, **30,11064** (1991).

103. T. Stams, S. K. Nair, T. Okuyama, A. Waheed, W. S. Sly, and D. W. Christianson, Proc. Nat. Acad. Sci. USA, **93,13589 (1996).**

104. J. Vidgren, L. A. Svensson, and A. Liljas, Nature, **368, 354 (1994).**

105. M. Pellegrini and D. F. Mierke, Biochemistry, **38,14775 (1999).**

106. S. Maignan, J. P. Guilloteau, S. Pouzieux, Y. M. Choi-Sledeski, M. R. Becker, S. I. Klein, W. R. Ewing, H. W. Pauls, A. P. Spada, and V. Mikol, J. Med. Chem., **43,3226 (2000).**

107. B. A. Katz, R. Mackman, C. Luong, K. Radika, A. Martelli, P. A. Sprengeler, J. Wang, H. Chan, and L. Wong, Chem. Biol., **7,299 (2000).**

108. K. Kamata, H. Kawamoto, T. Honma, T. Iwama, and S. H. Kim, Proc. Nat. Acad. Sci. USA, **95,6630 (1998).**

109. H. Brandstetter, A. Kuhne, W. Bode, R. Huber, W. Vondersaal, K. Wirthensohn, and R. A. Engh, *J.* Biol. Chem., **271, 29988 (1996).**

110. M. Adler, D. D. Davey, G. B. Phillips, S. H. Kim, J. Jancarik, G. Rumennik, D. L. Light, and M. Whitlow, Biochemistry, **39, 12534 (2000).**

111. A. Wei, R. Alexander, J. Duke, H. Ross, S. Rosenfeld, and C.-H. Chang, J. Mol. Biol., **283, 147 (1998).**

112. K. Padmanabhan, K. P. Padmanabhan, A. Tulinsky, C. H. Park, W. Bode, R. Huber, D. T. Blankenship, A. D. Cardin, and W. Kisiel, J. Mol. Biol., **232,947 (1993).**

113. M. G. Malkowski, P. D. Martin, J. C. Guzik, and B. F. P. Edwards, Protein Sci., **6, 1438 (1997).**

114. A. Vandelocht, W. Bode, R. Huber, B. F. Lebonniec, S. R. Stone, C. T. Esmon, and M. T. Stubbs, Embo J.,**16,2977 (1997).**

115. E. Zhang and A. Tulinsky, Biophys. Chem., **63, 185** (1997).

116. H. Nar, M. Bauer, A. Schmid, J. Stassen, W. Wienen, H. W. Priepke, I. K. Kauffmann, U. J. Ries, and N. H. Hauel, Structure, **9, 29 (2001).**

117. A. Zdanov, S. Wu, J. DiMaio, Y. Konishi, Y. Li, X. Wu, B. F. Edwards, P. D. Martin, and M. Cygler, Proteins, **17,252 (1993)**

118. N. Y. Chirgadze, D. J. Sall, S. L. Briggs, D. K. Clawson, M. Zhang, G. F. Smith, and R. W. Schevitz, Protein Sci., **9, 29 (2000).**

119. L. Tabernero, C. Y. Chang, S. Ohringer, W. F. Lau, E. J. Iwanowicz, W.-C. Han, T. C. Wang, S. M. Seiler, D. G. M. Roberts, and J. S. Sack, *J.* Mol. Biol., **246, 14 (1995).**

120. J. A. Malikayil, J. P. Burkhart, H. A. Schreuder, R. J. Broersmajunior, C. Tardif, L. W. Kutcheriii, S. Mehdi, G. L. Schatzman, B. Neises, and N. P. Peet, Biochemistry, **36,1034 (1997).**

121. P. C. Weber, S. L. Lee, F. A. Lewandowski, M. C. Schadt, C. W. Chang, and C. A. Kettner, Biochemistry, **34, 3750 (1995).**

122. J. A. Huntington and C. T. Esmon, In press.

123. R. Krishnan, A. Tulinsky, G. P. Vlasuk, D. Pearson, P. Vallar, P. Bergum, T. K. Brunck, and W. C. Ripka, Protein Sci., **5, 422 (1996).**

124. R. A. Engh, H. Brandstetter, G. Sucher, A. Eichinger, U. Baumann, W. Bode, R. Huber, T. Poll, R. Rudolph, and W. Vondersaal, Structure *(Lond)*, **4, 1353 (1996).**

125. A. Lombardi, G. Desimone, F. Nastri, S. Galdiero, R. Dellamorte, N. Staiano, C. Pedone, M. Bolognesi, and V. Pavone, Protein Sci., **8, 91 (1999).**

126. E. Guinto, S. Caccia, T. Rose, K. Futterer, G. Waksman, and E. Dicera, Proc. Nat. Acad. Sci. USA, **96,1852 (1999).**

127. B. E. Maryanoff, X. Qiu, K. P. Padmanabhan, A. Tulinsky, H. R. Almondjunior, P. Andrade-Gordon, M. N. Greco, J. A. Kauffman, K. C. Nicolaou, A. Liu, P. H. Brungs, and N. Fusetani, Proc. Nat. Acad. Sci. USA, **90, 8048 (1993).**

128. B. A. Katz, J. M. Clark, J. S. Finer-Moore, T. E. Jenkins, C. R. Johnson, M. J. Ross, C. Luong, W. R. Moore, and R. M. Stroud, Nature, **391, 608 (1998).**

129. J. H. Matthews, R. Krishnan, M. J. Costanzo, B. E. Maryanoff, and A. Tulinsky, Biophys. J., **71,2830 (1996).**

130. B. Bachand, M. Tarazi, Y. St-Denis, J. J. Edmunds, P. D. Winocour, L. Leblond, and M. A. Siddiqui, Bioorg. Med. Chem. Lett., **11, 287 (2001).**

131. J. J. Slon-Usakiewicz, J. Sivaraman, Y. Li, M. Cygler, and Y. Konishi, Biochemistry, **39,2384 (2000).**

132. R. Krishnan, E. Zhang, K. Hakansson, R. K. Arni, A. Tulinskym.S. Lim-Wilby, O. E. Levy, J. E. Semple, and T. K. Brunck, Biochemistry, **37,12094 (1998).**

133. R. Krishnan, I. Mochalkin, R. Arni, and A. Tulinsky, *Acta Crystallogr.*, Sect. D, **56, 294 (2000).**

134. I. Mochalkin and A. Tulinsky, *Acta Crystallogr.*, Sect. D, **55, 785 (1999).**

135. R. Bone, T. Lu, C. R. Illig, R. M. Soll, and J. C. Spurlino, **J. *Med. Chem.*, 41,2068 (1998).**

136. R. Krishnan, E. J. Sadler, and A. Tulinsky, *Acta Crystallogr.*, Sect. D, **56,406** (2000).

137. M. Nardini, A. Pesce, M. Rizzi, E. Casale, R. Ferraccioli, G. Balliano, P. Milla, P. Ascenzi, and M. Bolognesi, **J. *Mol. Biol.*, 258, 851 (1996).**

138. M. F. Malley, L. Tabernero, C. Y. Chang, S. L. Ohringer, D. G. Roberts, J. Das, and J. S. Sack, ***Protein Sci.*, 5,221 (1996).**

139. J. L. R. Steiner, M. Murakami, and A. Tulinsky, J. ***Am. Chem. Soc.*, 120,597 (1998).**

140. I. I. Mathews and A. Tulinsky, ***Acta Crystallogr. D Biol. Crystallogr.*, 51,550 (1995).**

141. I. I. Mathews, K. P. Padmanabhan, V. Ganesh, A. Tulinsky, M. Ishil, J. Chen, C. W. Turck, and S. R. Coughlin, ***Biochemistry*, 33, 3266 (1994).**

142. J. Vijayalakshmi, K. P. Padmanabhan, K. G. Mann, and A. Tulinsky, ***Protein Sci.*, 3, 2254 (1994).**

143. 1. I. Mathews, K. P. Padmanabhan, A. Tulinsky, and J. E. Sadler, ***Biochemistry*, 33,13547** (1994).

144. J. P. Priestle, J. Rahuel, H. Rink, M. Tones, and M. G. Gruetter, ***Protein Sci.*, 2, 1630 (1993).**

145. T. J. Rydel, A. Tulinsky, W. Bode, and R. Huber, **J. *Mol. Biol.*, 221, 583** (1991).

146. D. W. Banner and P. Hadvary, **J. *Biol. Chem.*, 266,20085 (1991).**

147. R. K. Arni, K. Padmanabhan, K. P. Padmanabhan, T-P. Wu, and A. Tulinsky, ***Biochemistry*, 32,4727 (1993)**

148. X. Qiu, K. P. Padmanabhan, V. E. Carperos, A. Tulinsky, T. Kline, J. M. Maraganore, and J. W. Fentonii, ***Biochemistry*, 31,11689** (1992).

149. C. Eigenbrot, D. Kirchhofer, M. S. Dennis, L. Santell, R. A. Lazarus, J. Stamos, and M. H. Ultsch, ***Structure*, 9, 627** (2001).

150. A. C. W. Pike, A. M. Brzozowski, S.M. Roberts, O. H. Olsen, and E. Persson, ***Proc. Nat. Acad. Sci. USA*, 96,8925 (1999).**

151. M. S. Dennis, C. Eigenbrot, N. J. Skelton, M. H. Ultsch, L. Santell, M. L. Dwyer, M. P. O'Connell, and R. A. Lazarus, ***Nature*, 404, 465 (2000).**

152. D. W. Banner, A. D'Arcy, C. Chene, F. K. Winkler, A. Guha, W. H. Konigsberg, Y. Nemerson, and D. Kirchhofer, ***Nature*, 380**, 41 **(1996).**

153. G. Kemball-Cook, D. J. D. Johnson, E. G. D. Tuddenham, and K. Harlos, ***J. Struct. Biol.*, 127,213 (1999).**

154. E. Zhang, R. Stcharles, and A. Tulinsky, **J. *Mol. Biol.*, 285,2089 (1999).**

155. K.-P. Hopfner, A. Lang, A. Karcher, K. Sichler, E. Kopetzkih. Brandstetter, R. Huber, W. Bode, and R. A. Engh, ***Structure (Lond)*, 7,989 (1999).**

156. H. Brandstetter, M. Bauer, R. Huber, P. Lollar, and W. Bode, ***Proc. Nat. Acad. Sci. USA*, 92,9796** (1995).

157. M. G. Malkowski, S. L. Ginell, W. L. Smith, and R. M. Garavito, ***Science*, 289, 1933 (2000).**

158. D. Picot, P. J. Loll, and R. M. Garavito, ***Nature*, 367, 243 (1994).**

159. P. J. Loll, D. Picot, and R. M. Garavito, ***Nat. Struct. Biol.*, 2,637** (1995).

160. P. J. Loll, C. T. Sharkey, S. J. O'Connor, C. M. Dooley, E. O'Brien, M. Devocelle, K. B. Nolan, B. S. Selinsky, and D. J. Fitzgerald, ***Mol. Pharmacol.*, 60, 1407** (2001).

161. E. D. Thuresson, M. G. Malkowski, K. M. Lakkides, C. J. Riekea.M. Mulichak, S. L. Ginell, R. M. Garavito, and W. L. Smith, ***J. Biol. Chem.*, 276, 10358 (2001).**

162. B. S. Selinsky, K. Gupta, C. T. Sharkey, and P. J. Loll, ***Biochemistry*, 40, 5172 (2001).**

163. P. J. Loll, D. Picot, O. Ekabo, and R. M. Garavito, ***Biochemistry*, 35, 7330** (1996).

164. J. R. Kiefer, J. L. Pawlitz, K. T. Moreland, R. A. Stegeman, J. K. Gierse, W. F. Hood, J. K. Gierse, A. M. Stevens, D. C. Goodwin, S. W. Rowlinson, L. J. Marnett, W. C. Stallings, and R. G. Kurumbail, ***Nature*, 405**, 97 **(2000).**

165. R. G. Kurumbail, A. M. Stevens, J. K. Gierse, J. J. Mcdonald, R. A. Stegeman, J. Y. Pak, D. Gildehaus, J. M. Miyashiro, T. D. Penning, K. Seibert, P. C. Isakson, and W. C. Stallings, ***Nature*, 384,644 (1996).**

166. Q. Zhao, S. Modi, G. Smith, M. Paine, P. D. Mcdonagh, C. R. Wolf, D. Tew, L. Y. Lian, G. C. Roberts, and H. P. Driessen, ***Protein Sci.***, 8, **298 (1999).**

167. M. Wang, D. L. Roberts, R. Paschke, T. M. Shea, B. S. Masters, and J. J. Kim, ***Proc. Nat. Acad. Sci. USA*, 94,8411** (1997).

168. P. A. Hubbard, A. L. Shen, R. Paschke, C. B. Kasper, and J. J. Kim, **J. *Biol. Chem.*, 276, 29163 (2001).**

169. A. Gangjee, A. P. Vidwans, A. Vasudevan, S. F. Queener, R. L. Kisliuk, V. Cody, R. Li, N. Gal-

itsky, J. R. Luft, and W. Pangborn, J. Med. Chem., **41,3426 (1998).**

170. V. Cody, N. Galitsky, J. R. Luft, W. Pangborn, R. L. Blakley, and A. Gangjee, J. Mol. Biol., **221,583 (1991).**

171. V. Cody, N. Galitsky, J. R. Luft, W. Pangborn, A. Rosowsky, and R. L. Blakley, Biochemistry, **36,13897 (1997).**

172. M. A. McTigue, J. F. Davies II, B. T. Kaufman, and J. Kraut, Biochemistry, **31,7264 (1992).**

173. M. A. McTigue, J. F. Davies II, B. T. Kaufman, and J. Kraut, Biochemistry, **32,6855 (1993).**

174. M. A. McTigue, J. F. Daviesii, B. T. Kaufman, N.-H. Xuong, and J. Kraut, In Press.

175. J. F. Davies, T. J. Delcamp, N. J. Prendergast, V. A. Ashford, J. H. Freisheim, and J. Kraut, Biochemistry, **29,9467 (1990).**

176. W. S. Lewis, V. Cody, N. Galitsky, J. R. Luft, W. Pangborn, S. K. Chunduru, H. T. Spencer, J. R. Appleman, and R. L. Blakley, J. Biol. Chem., **270,5057 (1995).**

177. J. F. Davies, D. A. Matthews, S. J. Oatley, B. T. Kaufman, N.-H. Xuong, and J. Kraut, In press.

178. C. Oefner, A. D'Arcy, and F. K. Winkler, Eur. J. Biochem., **174,377 (1988).**

179. S. Liu, E. A. Neidhardt, T. H. Grossman, T. Ocain, and J. Clardy, Structure *(Lond)*, **8, 25 (2000).**

180. I. C. Hampele, A. D'Arcy, G. E. Dale, D. Kostrewa, J. Nielsenc. Oefner, M. G. Page, H. J. Schonfeld, D. Stuber, and R. L. Then, J. Mol. Biol., **268, 21 (1997).**

181. P. Soultanas, M. S. Dillingham, S. S. Velankar, and D. B. Wigley, J. Mol. Biol., **290,137 (1999).**

182. M. R. Redinbo, J. J. Champoux, and W. G. Hol, Biochemistry, **39, 6832 (2000).**

183. L. Stewart, M. R. Redinbo, X. Qiu, W. G. Hol, and J. J. Champoux, Science, **279,1534 (1998).**

184. M. R. Redinbo, L. Stewart, P. Kuhn, J. J. Champoux, and W. G. Hol, Science, **279,1504 (1998).**

185. M. Ruff, M. Gangloff, S. Eiler, S. Duclaud, J. M. Wurtz, and M. Dino, In press.

186. J. W. R. Schwabe, L. Chapman, J. T. Finch, D. Rhodes, and D. Neuhaus, Structure *(Lond)*, **1, 187 (1993).**

187. D. M. Tanenbaum, Y. Wang, S. P. Williams, and P. B. Sigler, Proc. *Nat.* Acad. Sci. USA, **95, 5998 (1998).**

188. A. M. Brzozowski, A. C. W. Pike, Z. Dauter, R. E. Hubbard, T. Bonn, O. Engstrom, L. Ohman, G. L. Greene, J.-A. Gustaffson, and M. Carlquist, Nature, **389, 753 (1997).**

189. J. W. R. Schwabe, L. Chapman, J. T. Finch, and D. Rhodes, Cell, **75,567 (1993).**

190. A. K. Shiau, D. Barstad, P. M. Loria, L. Cheng, P. J. Kushner, D. A. Agard, and G. L. Greene, Cell, **95,927 (1998).**

191. J. W. R. Schwabe, L. Chapman, J. T. Finch, and D. Rhodes, Cell, **75, 567 (1993).**

192. A. K. Shiau, D. Barstad, P.M. Loria, L. Cheng, P. J. Kushner, D. A. Agard, and G. L. Greene, Cell, **95,927** (1998).

193. K. P. Wilson, M. M. Yamashita, M. D. Sintchak, S. H. Rotsteinm.A. Murcko, J. Boger, J. A. Thomson, M. J. Fitzgibbon, and M. A. Navia, Acta Crystallogr., Sect. D, **51, 511 (1995).**

194. J. Choi, J. Chen, S. L. Schreiber, and J. Clardy, Science, **273,239 (1996).**

195. J. Liang, J. Clardy, and J. Choi, Acta Crystallogr., Sect. D, **55, 736 (1999).**

196. J. Liang, J. Choi, and J. Clardy, Acta Crystallogr. D Biol. Crystallogr., **55, 736 (1999).**

197. S. W. Michnick, M. K. Rosen, T. J. Wandless, M. Karplus, and S. L. Schreiber, Science, **252, 836 (1991)**

198. C. T. Rollins, V. M. Rivera, D. N. Woolfson, T. Keenan, M. Hatada, S. E. Adams, L. J. Andrade, D. Yaeger, M. R. Vanschravendijk, D. A. Holt, M. Gilman, and T. Clackson, Proc. *Nat.* Acad. Sci. USA, **97, 7096** (2000).

199. T. Clackson, W. Yang, L. Rozamus, J. Amara, M. H. Hatada, C. T. Rollins, L. F. Stevenson, S. R. Magari, S. A. Wood, N. L. Courage, X. Lu, F. Cerasolijunior, M. Gilman, and D. Holt, Proc. *Nat.* Acad. Sci. USA, **95, 10437 (1998).**

200. P. Burkhard, P. Taylor, and M. D. Walkinshaw, J. Mol. Biol., **295,953 (2000).**

201. J. W. Becker, J. Rotonda, J. G. Cryan, M. Martin, W. H. Parsons, P. J. Sinclair, G. Wiederrecht, and F. Wong, J. Med. Chem., **42, 2798 (1999).**

202. C. Sich, S. Improta, D. J. Cowley, C. Guenet, J. P. Merly, M. Teufel, and V. Saudek, Eur. J. Biochem., **267, 5342 (2000).**

203. M. Huse, Y. G. Chen, J. Massague, and J. Kuriyan, Cell, **96,425 (1999).**

204. L. W. Schultz and J. Clardy, Bioorg. Med. Chem. *Lett.*, **8, 1 (1998).**

205. S. Itoh, M. T. Decenzo, D. J. Livingston, D. A. Pearlman, and M. A. Navia, Bioorg. Med. Chem. Lett., **5, 1983** (1995).

206. J. Liang, J. Choi, and J. Clardy, Acta Crystallogr., Sect. D, **55, 736** (1999).

207. C. C. S. Deivanayagam, M. Carson, A. Thotakura, S. V. L. Narayana, and C. S. Chodavarapu, Acta Crystallogr., Sect. D, **56, 266 (2000).**

208. D. A. Holt, J. I. Luengo, D. S. Yamashita, H.-J. Oh, A. L. Konialian, H.-K. Yen, L. W. Rozamus, M. Brandt, M. J. Bossard, M. A. Levy, D. S. Eggleston, T. J. Stout, J. Liang, L. W. Schultz, and J. Clardy, ***J. Am. Chem. Soc.*, 115, 9925 (1993).**

209. G. D. VanDuyne, R. F. Standaert, P. A. Karplus, S. L. Schreiber, and J. Clardy, ***Science*, 252,839 (1991).**

210. G. D. Vanduyne, R. F. Standaert, S. L. Schreiber, and J. Clardy, ***J. Am. Chem. Soc.*, 113, 7433 (1991).**

211. K. M. Fox, J. A. Dias, and P. Vanroey, ***Mol. Endocrinol.*, 15,378 (2001).**

212. P. Storici, G. Capitani, D. Debiase, M. Moser, R. A. John, J. N. Jansonius, and T. Schirmer, ***Biochemistry*, 38,8628 (1999).**

213. D. T. Gewirth and P. B. Sigler, ***Nat. Struct. Biol.*, 2,386 (1995).**

214. B. F. Luisi, W. X. Xu, Z. Otwinowski, L. P. Freedman, K. R. Yamamoto, and P. B. Sigler, ***Nature*, 352,497 (1991).**

215. N. Kunishima, Y. Shimada, Y. Tsuji, T. Sato, M. Yamamoto, T. Kumasaka, S. Nakanishi, H. Jingami, and K. Morikawa, ***Nature*, 407, 971 (2000).**

216. O. Epp, R. Ladenstein, and A. Wendel, ***Eur. J. Biochem.*, 133, 51 (1983)**

217. M. Aritomi, N. Kunishima, T. Okamoto, R. Kuroki, Y. Ota, and K. Morikawa, ***Nature*, 401, 713 (1999).**

218. B. Lovejoy, D. Cascio, and D. Eisenberg, ***J. Mol. Biol.*, 234, 640 (1993).**

219. T. Zink, A. Ross, K. Luers, C. Ciesler, R. Rudolph, and T. A. Hola, ***Biochemistry*, 33, 8453 (1994).**

220. C. P. Hill, T. D. Osslund, and D. Eisenberg, ***Proc. Natl. Acad. Sci. USA*, 90,5167 (1993).**

221. D. Rozwarski, K. Diederichs, R. Hecht, T. Boone, and P. A. Karplus, ***Proteins*, 26, 304 (1996).**

222. T. Clackson, M. H. Ultsch, J. A. Wells, and A.M. Devos, ***J. Mol. Biol.*, 277, 1111 (1998).**

223. S. Atwell, M. Ultsch, A. M. Devos, and J. A. Wells, ***Science*, 278, 1125 (1997).**

224. M. Sundstrom, T. Lundqvist, J. Rodin, L. B. Giebel, D. Milligan, and G. Norstedt, ***J. Biol. Chem.*, 271,32197 (1996).**

225. A. M. Devos, M. Ultsch, and A. A. Kossiakoff, ***Science*, 255, 306 (1992).**

226. Y. Hsiou, J. Ding, K. Das, A. D. Clarkjunior, S. H. Hughes, and E. Arnold, ***Structure*, 4,853 (1996).**

227. J. Ren, R. M. Esnouf, A. L. Hopkins, E. Y. Jones, I. Kirby, J. Keeling, C. K. Ross, B. A. Larder, D. I. Stuart, and D. K. Stammers, ***Proc. Nat. Acad. Sci. USA*, 95,9518 (1998).**

228. Y. Hsiou, J. Ding, K. Das, A. D. Clark, P. L. Boyer, P. Lewi, P. A. J. Janssen, J.-P. Kleim, M. Roesner, S. H. Hughes, and E. Arnold, ***J. Mol. Biol.*, 309,437 (2001).**

229. Y. Hsiou, K. Das, J. Ding, A. D. Clarkjunior, J. P. Kleim, M. Rosner, I. Winkler, G. Riess, S. H. Hughes, and E. Arnold, ***J. Mol. Biol.*, 284, 313 (1998).**

230. K. Das, J. Ding, Y. Hsiou, A. D. Clarkjunior, H. Moereels, L. Koyrnans, K. Andries, R. Pauwels, P. A. Janssen, P. L. Boyer, P. Clark, R. H. Smithjunior, M. B. Kroegersmith, C. J. Michejda, S. H. Hughes and E. Arnold, ***J. Mol. Biol.*, 264, 1085 (1996).**

231. M. Hogberg, C. Sahlberg, P. Engelhardt, R. Noreen, J. Kangasmetsa, N. G. Johansson, B. Oberg, L. Vrang, H. Zhang, B. L. Sahlberg, T. Unge, S. Lovgren, K. Fridborg, and K. Backbro, ***J. Med. Chem.*, 43,304 (2000).**

232. J. Lindberg, S. Sigurdsson, S. Lowgren, H. O. Andersson, C. Sahlberg, R. Noreen, K. Fridborg, H. Zhang, and T. Unge, ***Eur. J. Biochem.*, 269, 1670 (2002).**

233. J. Jaeger, T. Restle, and T. A. Steitz, ***Embo J.*, 17,4535 (1998).**

234. A. L. Hopkins, J. Ren, H. Tanaka, B. Baba, M. Okamato, D. I. Stuart, and D. K. Stammers, ***J. Med. Chem.*, 42,4500 (1999).**

235. J. Ding, K. Das, H. Yu, S. G. Sarafianos, A. D. Clarkjunior, A. Jacobo-Molina, C. Tantillo, S. H. Hughes, and E. Arnold, ***J. Mol. Biol.*, 284, 1095 (1998).**

236. S. G. Sarafianos, K. Das, C. Tantillo, A. D. Clarkjr., J. Ding, J. Whitcomb, P. L. Boyer, S. H. Hughes, and E. Arnold, ***Embo J.*, 20, 1449 (2001).**

237. D. W. Rodgers, S. J. Gamblin, B. A. Harris, S. Ray, J. S. Culp, B. Hellmig, D. J. Woolf, C. Debouck, and S. C. Harrison, ***Proc. Natl. Acad. Sci. USA*, 92,1222 (1995).**

238. J. Ding, K. Das, C. Tantillo, W. Zhang, A. D. Clarkjunior, S. Jessen, X. Lu, Y. Hsiou, A. Jacobo-Molina, K. Andries, R. Pauwels, H. Moereels, L. Koymans, P. A. J. Janssen, R. H. Smithjunior, M. K. Koepke, C. J. Michejda, S. H. Hughes, and E. Arnold, ***Structure (Lond)*, 3, 365 (1995).**

239. J. Ding, K. Das, H. Moereels, L. Koymans, K. Andries, P. A. J. Janssen, S. H. Hughes, and E. Arnold, ***Nat. Struct. Biol.*, 2, 407 (1995).**

240. J. Ren, J. Milton, K. L. Weaver, S. A. Short, D. I. Stuart, and D. K. Stammers, Structure, **8, 1089 (2000).**

241. J. Ren, C. Nichols, L. Bird, P. Chamberlain, K. Weaver, S. Short, D. I. Stuart, and D. K. Stammers, J. Mol. Biol., **312, 795 (2001).**

242. S. G. Sarafianos, K. Das, C. Tantillo, A. D. Clarkjr., J. Ding, J. Whitcomb, P. L. Boyer, S. H. Hughes, and E. Arnold, Proc. Nat. Acad. Sci. USA, **96,10027 (1999).**

243. J. Wang, S. J. Smerdon, J. Jager, L. A. Kohlstaedt, P. A. Rice, J. M. Friedman, T. A. Steitz, Proc. Natl. Acad. Sci. USA, **91, 7242 (1994).**

244. M. D. Sintchak, M. A. Fleming, O. Futer, S. A. Raybuck, S. P. Chambers, P. R. Caron, M. A. Murcko, and K. P. Wilson, Cell, **85,921** (1996).

245. T. D. Colby, K. Vanderveen, M. Strickler, G. D. Markham, and B. M. Goldstein, Proc. Nat. Acad. Sci. USA, **96, 3531** (1999).

246. L. G. Laajoki, G. L. Francis, J. C. Wallace, J. A. Carver, and M. A. Keniry, J. Biol. Chem., **275, 10009 (2000).**

247. A. Sato, S. Nishimura, T. ohkubo, Y. Kyogoku, S. Koyama, M. Kobayashi, T. Vasuda, and Y. Kobayashi, Znt. J. Pept. Protein Res., **41, 433 (1993).**

248. F. F. Vajdos, M. Ultsch, M. L. Schaffer, K. D. Deshayes, J. Liu, N. J. Skelton, and A. M. Devos, Biochemistry, **40, 11022** (2001).

249. E. Dewolf, R. Gill, S. Geddes, J. Pitts, A. Wollmer, and J. Grotzinger, Protein Sci., **5, 2193** (1996).

250. R. M. Cooke, T. S. Harvey, and I. D. Campbell, Biochemistry, **30, 5484 (1991).**

251. T. P. J. Garrett, N. M. Mckern, M. Lou, M. J. Frenkel, J. D. Bentley, G. O. Lovrecz, T. C. Elleman, L. J. Cosgrove, and C. W. Ward, Nature, **394,395** (1998).

252. K. Parang, J. H. Till, A. J. Ablooglu, R. A. Kohanski, S. R. Hubbard, and P. A. Cole, Nat. Struct. Biol., **8, 37 (2001).**

253. J. H. Till, A. J. Ablooglu, M. Frankel, S. M. Bishop, R. A. Kohanski, and S. R. Hubbard, J. Biol. Chem., **276, 10049** (2001).

254. S. R. Hubbard, Embo J., **16, 5572 (1997).**

255. S. R. Hubbard, L. Wei, L. Ellis, and W. A. Hendrickson, Nature, **372,746 (1994).**

256. A. M. Torres, B. E. Forbes, S. E. Aplin, J. C. Wallace, G. L. Francis, and R. S. Norton, *J.* Mol. Biol., **248,385** (1995).

257. E. T. Baldwin, R. W. Sarver, G. L. Bryantjunior, K. A. Currym.B. Fairbanks, B. C. Finzel, R. L. Garlick, R. L. Heinrikson, N. C. Horton, L. L. Kelley, A. M. Mildner, J. B. Moon, J. E. Mott, V. T. Mutchler, C. S. Tomich, K. D. Watenpaugh, and V. H. Wiley, Structure (Land), **6,923 (1998).**

258. J. O. Lee, L. A. Bankston, M. A. Arnaout, and R. C. Liddington, Structure (Land), **3, 1333 (1995).**

259. J. O. Lee, P. Rieu, M. A. Arnaout, and R. Liddington, Cell, **80,631** (1995).

260. J. Bella, P. R. Kolatkar, C. W. Marlor, J. M. Greve, and M. G. Rossmann, Proc. Nat. Acad. Sci. USA, **95,4140 (1998).**

261. J. M. Casasnovas, T. Stehle, J. H. Liu, J. H. Wang, and T. A. Springer, Proc. Nat. Acad. *Sci.* USA, **95,4134 (1998).**

262. P. R. Kolatkar, J. Bella, N. H. Olson, C. Bator, T. S. Baker, and M. G. Rossmann, Embo J., **18, 6249 (1999).**

263. W. Klaus, B. Gsell, A. M. Labhardt, B. Wipf, and H. Senn, *J.* Mol. Biol., **274,661 (1997).**

264. R. Radhakrishnan, L. J. Walter, A. Hruza, P. Reichert, P. P. Trotta, T. L. Nagabhushan, and M. R. Walter, Structure, **4, 1453 (1996)**

265. D. J. Thiel, M.-H. Ledu, R. L. Walter, A. D'Arcy, C. Chene, M. Fountoulakis, G. Garotta, F. K. Winkler, and S. E. Ealick, Structure Fold Des., **8,927** (2000).

266. M. Randal and A. A. Kossiakoff, Structure, **9, 155 (2001).**

267. A. Landar, B. Curry, M. H. Parker, R. Digiacomo, S. R. Indelicato, T. L. Nagabhushan, G. Rizzi, and M. R. Walter, J. Mol. Biol., **299, 169** (2000).

268. S. E. Ealick, W. J. Cook, S. Vijay-Kumar, M. Carson, T. L. Nagabhushan, P. P. Trotta, and C. E. Bugg, Science, **252,698** (1991).

269. B. J. Graves, M. H. Hatada, W. A. Hendrickson, J. K. Miller, V. S. Madison, and Y. Satow, Biochemistry, **29,2679 (1990).**

270. G. P. A. Vigers, D. J. Dripps, C. K. Edwards, and B. J. Brandhuber, Structure, **275, 36927 (2000).**

271. H. Schreuder, C. Tardif, S. Trump-Kallmeyer, A. Soffientini, E. Sarubbi, A. Akeson, T. Bowlin, S. Yanofsky, and R. W. Barrett, Nature, **386, 194** (1997).

272. G. P. Vigers, L. J. Anderson, P. Caffes, and B. J. Brandhuber, Nature, **386, 190 (1997).**

273. A. Zdanov, C. Schalk-Hihi, S. Menon, K. W. Moore and A. Wlodawer, J. Mol. Biol., **268,460 (1997).**

274. A. Zdanov, C. Schalk-Hihi, and A. Wlodawer, Protein Sci., **5, 1955 (1996).**

275. A. Zdanov, C. Schalk-Hihi, A. Gustchina, M. Tsang, J. Weatherbee, and A. Wlodawer, Structure, **3,591 (1995).**

276. K. Josephson, N. J. Logsdon, and M. R. Walter, Immunity, **15, 35 (2001).**

277. M. R. Walter and T. L. Nagabhushan, Biochemistry, **34, 12118 (1995).**

278. C. Yoon, S. C. Johnston, J. Tang, M. Stahl, J. F. Tobin, and W. S. Somers, Embo **J., 19, 3530 (2000).**

279. E. Z. Eisenmesser, D. A. Horita, A. S. Altieri, and R. A. Byrd, *J.* Mol. Biol., **310,231 (2001).**

280. H. R. Mott, B. S. Baines, R. M. Hall, R. M. Cooke, P. C. Driscoll, M. P. Weir, and I. D. Campbell, **J.** Mol. Biol., **248, 979 (1995).**

281. D. B. McKay, Science, **257,412 (1992).**

282. Y. Feng, B. K. Klein, and C. A. Mcwherter, **J.** Mol. Biol., **259,524 (1996).**

283. T. Mueller, **F.** Oehlenschlaeger, and M. Buehner, **J.** Mol. Biol., **247, 360 (1955).**

284. T. Hage, W. Sebald, and P. Reinemer, Cell, **97, 271 (1999).**

285. M. Hulsmeyer, C. Scheufler, and M. K. Dreyer, *Acta* Crystallogr., Sect. D, **57, 1334 (2001).**

286. C. Redfield, L. J. Smith, J. Boyd, G. M. P. Lawrence, R. G. Edwards, C. J. Gershater, R. A. G. Smith, and C. M. Dobson, *J.* Mol. Biol., **238**, 23 (1994).

287. R. Powers, D. S. Garrett, C. J. March, E. A. Frieden, A. M. Gronenborn, and G. M. Clore, Science, **256, 1673 (1992).**

288. T. Mueller, T. Dieckmann, W. Sebald, and H. Oschkinat, *J.* Mol. *Biol.***237**, 423 **(1994)**

289. T. Mueller, T. Dieckmann, W. Sebald, and H. Oschkinat, *J.* Mol. Biol., **237,423 (1994).**

290. L. J. Smith, C. Redfield, J. Boyd, G. M. P. Lawrence, R. G. Edwards, R. A. G. Smith, and C. M. Dobson, *J.* Mol. Biol., **224,899 (1992).**

291. M. R. Walter, W. J. Cook, B. G. Zhao, R. Cameronjunior, S. E. Ealick, R. L. Walterjunior, P. Reichert, T. L. Nagabhushan, P. P. Trotta, and C. E. Bugg, *J.* Biol. Chem., **267,20371 (1992).**

292. A. Wlodawer, A. Pavlovsky, and A. Gustchina, Febs ***Lett.*, 309**, 59 **(1992).**

293. R. Powers, D. S. Garrett, C. J. March, E. A. Frieden, A. M. Gronenborn, and G. M. Clore, Biochemistry, **32,6744 (1993).**

294. M. V. Milburn, A. M. Hassell, M. H. Lambert, S. R. Jordan, A. E. I. Proudfoot, P. Graber, and T. N. C. Wells, Nature, **363, 172 (1993).**

295. G. Y. Xu, H. A. Yu, J. Hong, M. Stahl, T. Mcdonagh, L. E. Kay, and D. A. Cumming, *J.* Mol. Biol., **268,468 (1997).**

296. W. Somers, M. Stahl, and J. S. Seehra, Embo *J.*, **16, 989 (1997).**

297. K. Rajarathnam, I. Clark-Lewis, and B. D. Sykes, Biochemistry, **34, 12983 (1995).**

298. C. Eigenbrot, H. B. Lowman, L. Chee, and D. R. Artis, Proteins, **27, 556 (1997).**

299. N. J. Skelton, C. Quan, D. Reilly, and H. Lowman, Structure Fold Des., **7, 157 (1999).**

300. N. Gerber, H. Lowman, D. R. Artis, and C. Eigenbrot, Proteins: Struct., Funct., Genet., **38,361** (2000).

301. H. Sticht, M. Auer, B. Schmitt, J. Besemer, M. Horcher, T. Kirsch, J. D. Lindley, and P. Roesch, Eur. J. Biochem., **235, 26 (1996).**

302. E. T. Baldwin, I. T. Weber, R. St. Charles, J.-C. Xuan, E. Appella, M. Yamada, K. Matsushima, B. F. P. Edwards, G. M. Clore, A. M. Gronenborn, and A. Wlodawer, Proc. *Nat. Acad.* Sci. USA, **88,502 (1991).**

303. G. M. Clore, E. Appella, M. Yamada, K. Matsushima, and A. M. Gronenborn, Biochemistry, **29,1689 (1990).**

304. M. M. G. M. Thunnissen, P. N. Nordlund, and J. Z. Haeggstrom, *Nat.* Struct. Biol., **8, 131 (2001).**

305. X. Weng, H. Luecke, I. S. Song, D. S. Kang, S-H. Kim, and R. Huber, Protein Sci., **2, 448 (1993).**

306. A. Rosengarth, V. Gerke, and H. Luecke, *J.* Mol. Biol., **306,489 (2001).**

307. M. Tegoni, S. Spinelli, M. Verhoeyen, P. Davis, and C. Carnbilla, *J.* Mol. Biol., **289, 1375 (1999).**

308. H. Wu, J. W. Lustbader, Y. Liu, R. E. Canfield, and W. A. Hendrickson, Structure, **2, 545 (1994).**

309. A. J. Lapthorn, D. C. Harris, A. Littlejohn, J. W. Lustbader, R. E. Canfield, K. J. Machin, F. J. Morgan, and N. W. Isaacs, Nature, **369, 455 (1994).**

310. A. Bohm, J. Pandit, J. Jancarik, R. Halenbeck, K. Koths, and S.-H. Kim, Science, **258, 1358 (1992).**

311. M. J. Jedrzejas, S. Singh, W. J. Brouillette, W. G. Laver, G. M. Air, and M. Luo, *J.* Mol. Biol., **267,584** (1997).

312. N. R. Taylor, A. Cleasby, O. Singh, T. Skarzynski, A. J. Wonacott, P. W. Smith, S. L. Sollis, P. D. Howes, P. C. Cherry, R. Bethell, P. Colman, and J. Varghese, J. Med. Chem., **41,798 (1998).**

313. M. J. Jedrzejas, S. Singh, W. J. Brouillette, W. G. Laver, G. M. Air, and M. Luo, Biochemistry, **34, 3144 (1995).**

314. W. P. Burmeister, B. Henrissat, C. Bosso, S. Cusack, and R. W. H. Ruigrok, ***Structure,* 1, 19 (1993).**

315. J. B. Finley, V. R. Atigadda, F. Duarte, J. J. Zhao, W. J. Brouillette, G. M. Air, and M. Luo, **J. *Mol. Biol.*, 293, 1107 (1999).**

316. C. L. White, M. N. Janakiraman, W. G. Laver, C. Philippon, A. Vasella, G. M. Air, and M. Luo, **J. *Mol. Biol.*, 245,623** (1995).

317. W. P. Burmeister, R. W. H. Ruigrok, and S. Cusack, ***Embo* J., 11, 49** (1992).

318. S. A. Monks, G. Karagianis, G. J. Howlett, and R. S. Norton, ***J. Biomol. NMR*, 8,379 (1996).**

319. R. Bader, A. Bettio, A. G. Beck-Sickinger, and O. Zerbe, **J. *Mol. Biol.*, 305,307 (2001).**

320. C. Cabrele, M. Langer, R. Bader, H. A. Wieland, H. N. Doods, O. Zerbe, and A. G. Beck-Sickinger, **J. *Biol. Chem.*, 275,36043 (2000).**

321. G. Seidel, W. Schaefer, A. Esswein, E. Hofmann, and P. Roesch, In Press.

322. Z. Chen, P. Xu, J.-R. Barbier, G. Willick, and F. Ni, ***Biochemistry*, 39, 12766 (2000).**

323. U. C. Mam, K. Adermann, P. Bayer, W.-G. Forssmann, and P. Roesc, ***Biochem. Biophys. Res. Comm.*, 267,213 (2000).**

324. L. Jin, S. L. Briggs, S. Chandrasekhar, N. Y. Chirgadze, D. K. Clawson, R. W. Schevitz, D. L. Smiley, A. H. Tashjian, and F. Zhang, **J. *Biol. Chem.*, 275,27238 (2000).**

325. U. C. Mam, S. Austermann, P. Bayer, K. Adermann, A. Ejcharth. Sticht, S. Walter, F.-X. Schmid, R. Jaenicke, W.-G. Forssmann and P. Roesch, **J. *Biol. Chem.*, 270, 15194 (1995).**

326. U. C. Marx, *Strukturen Verschiedener Parathormonfragmente in Loesung*, University of Bayreuth Thesis, Bayreuth, **1996.**

327. G. Y. Xu, T. Mcdonagh, H. A. Yu, E. A. Nalefski, J. D. Clark, and D. A. Cumming, **J. *Mol. Biol.*, 280,485 (1998).**

328. O. Perisic, S. Fong, D. E. Lynch, M. Bycroft, and R. L. Williams, **J. *Biol. Chem.*, 273, 1596 (1998).**

329. A. Dessen, J. Tang, H. Schmidt, M. Stahl, J. D. Clark, J. Seehra, and W. S. Somers, ***Cell*, 97, 349 (1999).**

330. A. Kreusch, P. J. Pfaffinger, C. F. Stevens, and S. Choe, ***Nature*, 392,945 (1998).**

331. S. J. Cushman, M. H. Nanao, A. W. Jahng, D. Derubeis, S. Choe, and P. J. Pfaffinger, ***Nat. Struct. Biol.*, 7,403 (2000).**

332. K. A. Bixby, M. H. Nanao, N. V. Shen, A. Kreusch, H. Bellamy, P. J. Pfaffinger, and S. Choe, ***Nat. Struct. Biol.*, 6**, 38 (1998).

333. J. M. Gulbis, M. Zhou, S. Mann, and R. Mackinnon, ***Science*, 289, 123 (2000).**

334. D. L. Minorjr., Y.-F. Lin, B. C. Mobley, A. Avelar, Y. N. Janl.Y. Jan, and J. M. Berger, ***Cell*, 102, 657 (2000).**

335. J. L. Oberfield, J. L. Collins, C. P. Holmes, D. M. Goreham, J. P. Cooper, J. E. Cobb, J. M. Lenhard, E. A. Hull-Ryde, C. P. Mohr, S. G. Blanchard, D. J. Parks, L. B. Moore, J. M. Lehmann, K. Plunket, A. B. Miller, M. V. Milburn, S. A. Kliewer, and T. M. Wilson, ***Proc. Nat. Acad. Sci. USA*, 96,6102 (1999).**

336. R. T. Nolte, G. B. Wisely, S. Westin, J. E. Cobb, M. H. Lambert, R. Kurokawa, M. G. Rosenfeld, T. M. Willson, C. K. Glass, and M. V. Milburn, ***Nature*, 395, 137 (1998).**

337. R. T. Gampejr., V. G. Montana, M. H. Lambert, A. B. Miller, R. K. Bledsoe, M. V. Milburn, S. A. Kliewer, T. M. Willson, and H. E. Xu, ***Mol. Cell*, 5, 545** (2000).

338. J. Uppenberg, C. Svensson, M. Jaki, G. Bertilsson, L. Jendeberg, and A. Berkenstam, **J. *Biol. Chem.*, 273,31108** (1998).

339. S. P. Williams and P. B. Sigler, ***Nature*, 393, 392 (1998).**

340. W. Somers, M. Ultsch, A. M. Devos, and A. A. Kossiakoff, ***Nature*, 372,478 (1994).**

341. P. A. Elkins, H. W. Christinger, Y. Sandowski, E. Sakal, A. Gertler, A. M. Devos, and A. A. Kossiakoff, ***Nat. Struct. Biol.*, 7, 808** (2000).

342. F. Rastinejad, T. Wagner, Q. Zhao, and S. Khorasanizadeh, ***Embo* J., 19, 1045 (2000).**

343. B. P. Klaholz, A. Mitschler, M. Belema, C. Zusi, and D. Moras, ***Proc. Nat. Acad. Sci. USA*, 97, 6322 (2000).**

344. J. P. Renaud, N. Rochel, M. Ruff, V. Vivat, P. Chambon, H. Gronemeyer, and D. Moras, ***Nature*, 378, 681 (1995).**

345. B. P. Klaholz, J. P. Renaud, A. Mitschler, C. Zusi, P. Chambon, H. Gronemeyer, and D. Moras, *Nat. Struct. Biol.*, 5, 199 (1998).**

346. W. Bourguet, V. Vivat, J. M. Wurtz, P. Chambon, H. Gronemeyer, and D. Moras, ***Mol. Cell*, 5, 289 (2000).**

347. B. P. Klaholz, A. Mitschler, and D. Moras, *J. Mol. Biol.*, 302, 155** (2000).

348. R. M. A. Knegtel, M. Katahira, J. G. Schilthuis, A. M. J. J. Bonvin, R. Boelens, D. Eib, P. T. Vandersaag, and R. Kaptein, **J. *Biomol. NMR*, 3, 1** (1993).

349. W. Bourguet, M. Ruff, P. Chambon, H. Gronemeyer, and D. Moras, ***Nature*, 375,377** (1995).

350. F. Rastinejad, T. Perlmann, R. M. Evans, and P. B. Sigler, ***Nature*, 375, 203** (1995).

351. S. M. Holmbeck, M. P. Foster, D. R. Casimiro, D. S. Sem, H. J. Dyson, and P. E. Wright, J. Mol. Biol., **281,271** (1998).

352. R. T. Gampejr., V. G. Montana, M. H. Lambert, G. B. Wisely, M. V. Milburn, and H. E. Xu, Genes Dev., **14, 2229 (2000).**

353. P. F. Egea, A. Mitschler, N. Rochel, M. Ruff, P. Chambon, and D. Moras, Embo J., **19, 2592 (2000).**

354. Q. Zhao, S. A. Chasse, S. Devarakonda, M. L. Sierk, and B. A. Rastinejad, J. Mol. Biol., **296, 509 (2000).**

355. D. R. Hall, J. M. Hadden, G. A. Leonard, S. Bailey, M. Neu, M. Winn, and P. F. Lindley, *Acta* Crystallogr., Sect. D, **58, 70 (2002).**

356. Z. Zhang, R. Zhang, A. Joachimiak, J. Schlessinger, and X. Kong, Proc. Nat. Acad. Sci. USA, **97, 7732 (2000).**

357. X. Jiang, O. Gurel, E. A. Mendiaz, G. W. Stearns, C. L. Clogston, H. S. Lu, T. D. Osslund, R. S. Syed, K. E. Langley, and W. A. Hendrickson, Embo J., **19,3192** (2000).

358. J. N. Charnpness, M. S. Bennett, F. Wien, R. Visse, W. C. Summers, P. Herdewijn, E. Declercq, T. Ostrowski, R. L. Jarvest, and M. R. Sanderson, Proteins, **32, 350** (1998).

359. M. S. Bennett, F. Wien, J. N. Champness, T. Batuwangala, T. Rutherford, W. C. Summers, H. Sun, G. Wright, and M. R. Sanderson, Febs Lett., **443, 121 (1999).**

360. J. N. Champness, M. S. Bennett, F. Wien, R. Visse, W. C. Summers, P. Herdewijn, E. Declerq, T. Ostrowski, R. L. Jarvest, and M. R. Sanderson, Proteins Struct. Funct. Genet., **32, 350 (1998).**

361. K. Wild, T. Bohner, G. Folkers, and G. E. Schulz, Protein Sci., **6,2097 (1997).**

362. J. Vogt, R. Perozzo, A. Pautsch, A. Prota, P. Schelling, B. Pilger, G. Folkers, L. Scapozza, and G. E. Schulz, Proteins Struct. Funct. Genet., **41, 545 (2000).**

363. C. Wurth, U. Kessler, J. Vogt, G. E. Schulz, G. Folkers, and L. Scapozza, Protein Sci., **10, 63 (2001).**

364. A. Prota, J. Vogt, B. Pilger, R. Perozzo, C. Wurth, V. Marquez, P. Russ, G. E. Schulz, G. Folkers, and L. Scapozza, Biochemistry, **39, 9597 (2000).**

365. J. H. Naismith, T. Q. Devine, B. Brandhuber, and S. R. Sprang, J. Biol. Chem., **270, 13303 (1995).**

366. J. H. Naismith, T. Q. Devine, H. Khono, and S. R. Sprang, Structure, **4, 1251 (1996).**

367. D. W. Banner, A. D'Arcy, W. Janes, R. Gentz, H-J. Schoenfeld, C. Broger, H. Loetscher, and W. Lesslauer, Cell, **73,431 (1993).**

368. G. Tocchini-Valentini, N. Rochel, J. M. Wurtz, A. Mitschler, and D. Moras, Proc. Nat. Acad. Sci. USA, **98, 5491 (2001).**

369. N. Rochel, J. M. Wurtz, A. Mitschler, B. Klaholz, and D. Moras, Mol. Cell, **5, 173** (2000).

370. S. Vos, R. J. Parry, M. R. Burns, J. Dejersey, and J. L. Martin, J. Mol. Biol., **282,875** (1998).

371. S. Vos and J. De Jersey, Biochemistry, **36,4125** (1997).

372. S. Vos, R. J. Parry, M. R. Burns, J. Dejersey, and J. L. Martin, J. Mol. Biol., **282,875** (1998).

373. H. A. Lewis, Structure *(Camb)*, 9, **527–537 (2001).**

374. C. Freiberg, Drug Discov. Today, **6, S72–S80 (2001).**

375. J. M. Sauder, J. W. Arthur, and R. L. Dunbrack, Proteins, **40, 6–22 (2000).**

376. S. W. Muchmore, Nature, **381,335–341 (1996).**

377. K. A. Denessiouk, A. I. Denesyuk, J. V. Lehtonen, T. Korpela, and M. S. Johnson, Proteins, **35, 250–261 (1999).**

378. M. Teplova, Protein Sci., **9,2557–2566 (2000).**

379. T. J. Boggon, W. S. Shan, S. Santagata, S. C. Myers, and L. Shapiro, Science, **286, 2119–2125 (1999).**

380. P. W. Kleyn, Cell, **85,281–290** (1996).

381. K. Noben-Trauth, J. K. Naggert, M. A. North, and P. M. Nishina, Nature, **380, 534–538 (1996).**

382. S. Santagata, Science, **292,2041–2050 (2001).**

383. W. Eisenreich, Chem. Biol., **5, R221–R233 (1998).**

384. R. Sanchez, *Nat.* Struct. Biol., **7, 986–990 (2000).**

385. A. Sali, Nat. Struct. Biol., **5,1029–1032** (1998).

386. A. Fiser, R. K. Do, and A. Sali, Protein Sci., **9, 1753–1773 (2000).**

387. K. T. Simons, C. Strauss, and D. Baker, J. Mol. Biol., **306, 1191–1199 (2001).**

388. R. Sanchez, Nucleic Acids Res., **28, 250–253 (2000).**