CHAPTER SEVEN

Docking and Scoring Functions1 Virtual Screening

CHRISTOPH SOTRIFFER
GERHARD KLEBE
University of Marburg
Department of Pharmaceutical Chemistry
Marburg, Germany

MARTIN STAHL
HANS JOACHIM BÖHM
Discovery Technologies
F. Hoffmann-La Roche AG
Basel, Switzerland

Contents

1 Introduction, 282
2 General Concepts and Physical Background, 284
  2.1 Protein–Ligand Interactions and the Physical Basis of Biomolecular Recognition, 284
  2.2 Docking, Scoring, and Virtual Screening: The Basic Concepts, 289
3 Docking, 290
  3.1 General Concepts to Address the Docking Problem, 291
    3.1.1 Representation of the Macromolecular Receptor, 291
    3.1.2 Ligand Handling, 293
    3.1.3 Strategies for Searching the Configuration and Conformation Space, 294
      3.1.3.1 Geometric/Combinatorial Search Strategies, 295
      3.1.3.2 Energy Driven/Stochastic Procedures, 296
  3.2 Special Aspects of Docking, 300
    3.2.1 Protein Flexibility, 300
    3.2.2 Water Molecules, 302
    3.2.3 Assessment of Docking Methods, 303
    3.2.4 Docking and QSAR, 304
    3.2.5 Docking and Homology Modeling, 305
4 Scoring Functions, 306
  4.1 Description of Scoring Functions for Protein–Ligand Interactions, 306

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.

4.1.1 Force Field–Based Methods, **306**
4.1.2 Empirical Scoring Functions, **308**
4.1.3 Knowledge-Based Methods, **310**
4.2 Critical Assessment of Current Scoring Functions, **312**
4.2.1 Influence of the Training Data, **312**
4.2.2 Molecular Size, **312**
4.2.3 Penalty Terms, **313**
4.2.4 Specific Attractive Interactions, **313**
4.2.5 Water Structure and Protonation State, **313**
4.2.6 Performance in Structure Prediction and **Rank** Ordering of Related Ligands, **314**
5 Virtual Screening, **315**
5.1 Combinatorial Docking, **317**
5.2 Seeding Experiments to Assess Docking and Scoring in Virtual Screening, **318**
5.3 Hydrogen Bonding versus Hydrophobic Interactions, **319**
5.4 Finding Weak Inhibitors, **319**
5.5 Consensus Scoring, **319**
5.6 Successful Identification of Novel Leads through Virtual Screening, **320**
6 Outlook, **321**
7 Acknowledgments, **322**

1 INTRODUCTION

The action of drug molecules and the function of protein targets are governed by principles of molecular recognition. Binding events between ligands and their receptors in biological systems form the basis of physiological activity and pharmacological effects of chemical compounds. Accordingly, the rational development of new drugs requires an understanding of molecular recognition in terms of both structure and energetics (1). With respect to practical applications, it requires tools that are based on such knowledge of mutual recognition between molecular structures. Docking and virtual screening are computational tools to investigate the binding between macromolecular targets and potential ligands. They constitute an essential part of structure-based drug design, the area of medicinal chemistry that harnesses structural information for the purpose of drug discovery.

Structure-based design has become an integral part of medicinal chemistry. Although the knowledge about molecular recognition and its foundation on structural principles is still far from being complete, it has already fueled significant advances and contributed to many success stories in drug discovery (2). Convincing evidence has been accumulated for a large number of targets that the protein three-dimensional (3D) structure can be used to design small molecule ligands binding tightly to the protein. Several marketed compounds can indeed be attributed to successful structure-based design (3–6), as summarized in a number of recent reviews (2, 7–9).

Structure-based drug design is an iterative process (2, 10). It requires as the starting point the crystal structure or a reliable homology model of the target protein, preferentially complexed with a ligand. The first step of the process is a detailed analysis of the binding site and a compilation of all aspects possibly responsible for binding affinity and selectivity. These data are then used to generate new ideas how to improve existing ligands or to develop alternative molecular frameworks. Computational methods and molecular modeling play an essential role in this phase of hypothesis generation. They help to exploit information about the binding site geometry by constructing new molecules de *novo*, by analyzing known molecules with respect to their affinity and binding geometry, or by searching compound libraries for potential hits to suggest new leads. Discovered hits that are commercially available or synthetically accessible are then experimentally tested and their binding properties examined by biochemical, crystallographic, and spectroscopic methods. The 3D structure of new complexes together with the acquired activity data are subsequently used to start a new cycle of ligand design to improve the hypotheses stated in the previous round.

Since the introduction of computational structure-based design techniques into the drug discovery process in the early 1980s, the impact of these methods has significantly changed. Initially, computational tools were

often applied a ***posteriori*** to rationalize and understand the binding and structure-activity relationships in a series of inhibitors and to assist in the manual design of individual compounds. Guided by the creativity of the designer, a novel putative ligand was constructed using computer graphics. Molecular mechanics calculations were performed on the produced protein–ligand complex to assess the properties of the generated ligand in terms of a geometric and energetics analysis. A ligand was assumed to bind with high affinity if satisfactory complementarity in shape and surface properties between the protein and the ligand could be detected.

It has been realized, however, that the design of a single, synthetically accessible, active compound is a larger challenge than anticipated. Many phenomena of molecular recognition are not yet fully understood, nor are current modeling tools able to reflect and accordingly predict them with sufficient reliability. Most important, a fast and accurate computational prediction of binding affinities for new inhibitor candidates is still difficult to obtain. Although the existing tools do certainly not allow the medicinal chemist to design the one perfect ligand, they can help to *enrich* sets of molecules with more active ones, even though the known deficiencies of the methods can still lead to significant rates of both false positives and false negatives. A more moderate goal of current molecular design is thus to improve the hit rates of molecules suggested for biological assaying compared to a mere random compound selection or testing of nontargeted compound libraries. This implies that structure-based design approaches now focus on the processing of large numbers of molecules, arranged in so-called virtual libraries. These can be composed of either existing chemical substances (such as, for example, compound collections of a pharmaceutical company) or hypothetical new molecules that could be synthesized by combinatorial chemistry. The task is then to filter these large libraries by eliminating the majority of molecules that is rather unlikely to bind and by prioritizing the remaining ones. As recent experience shows, this strategy can be successful: several publications have reported quite impressive enrichments of active compounds (11–16).

The change of focus from single molecule to compound library design in modern structure-based drug discovery is also a consequence of major technological advances that have dramatically enhanced the data throughput in a variety of fields:

1. Progress in gene technology, protein chemistry, and structure determination techniques have resulted in a tremendous increase in **protein structure information.** The number of publicly available 3D protein structures continues to grow exponentially, with further acceleration expected from the current initiatives of structural genomics (17). As a consequence, more and more design projects are based on structural information, and structure-based ligand design has become routine at all major pharmaceutical companies. On the other hand, the growing amount of structural knowledge also calls for automated methods that make this new wealth of data accessible and available.
2. Automation and miniaturization have led to the development of **high-throughput screening (HTS),** which is now a well-established process for large-scale biological testing. Libraries of several hundred thousand compounds are routinely screened against new targets, frequently on a time scale of less than 1 month.
3. The characteristics of synthetic chemistry have significantly changed with the introduction of **combinatorial and parallel chemistry** techniques. The trend continues to move away from the synthesis of individual compounds toward the generation of compound libraries, whose members are accessible through the same type of chemical reaction but different building reagents.
4. Massive data processing and computational tasks formerly requiring expensive supercomputers have become generally feasible by advances in **PC cluster com-**

puting, offering supercomputing power at unprecedented performance-to-price ratios.

To provide competitive advantage, structure–based design tools must be fast enough to process thousands of compounds per day using affordable computing resources. Several algorithms have been developed that allow virtual ligand screening based on high-throughput flexible docking (18, 19). More sophisticated methods may then be applied at a later stage of refinement, where a smaller set of compounds, containing the most promising hits, is subjected to a more detailed analysis. Essential elements of all these docking tools are scoring functions that translate computationally generated protein-ligand binding geometries into estimates of affinity.

Docking as a computational tool of structure-based drug design to predict protein–ligand interaction geometries and binding affinities is the subject of this chapter. Fundamental aspects of the docking process, the scoring of protein–ligand interactions, and the application of docking to virtual ligand screening are discussed. At first, a discussion of the underlying physical principles determining protein-ligand recognition is given (Section 2.1), followed by a description of the general concepts of docking, scoring, and virtual screening (Section 2.2). Subsequently, the current approaches to the docking problem are presented (Section 3.1), focusing on the search methods (Section 3.1.3) and the approaches used to represent protein and ligand structures in an efficient way (Sections 3.1.1–3.1.2). In addition, a number of special aspects is discussed (Section 3.2), including, for example, the issues of protein flexibility (Section 3.2.1) or the consideration of water molecules in the context of docking (Section 3.2.2). This is followed by a section on scoring functions used for docking. Three major classes of scoring functions are presented (Section 4.1) and subjected to critical assessment (Section 4.2). A final section is dedicated to virtual screening, illustrating general strategies (Section 5), special problems (Sections 5.1–5.5), and representative applications (Section 5.6).

Although the goal of this chapter is to highlight the most important aspects of docking in the context of structure-based drug design, a comprehensive discussion of all aspects of current docking methodologies would be beyond its scope. Protein-protein or protein-DNA docking (20–24), as well as the docking of small ligands to DNA or RNA as targets (25), are not explicitly covered. The focus will be on the docking of small molecules to protein binding sites, with emphasis on automated procedures. Interactive docking tools, as provided by some modeling software packages, or fragment-based *de* novo design methods are not discussed. Information about the latter is available through other reviews (26, 27) (see also Ref. 28 for a list of de novo design algorithms). Virtual screening is discussed only in the context of docking; database screening techniques based on molecular similarity or pharmacophore models are not considered (29, 30). As a final introductory remark, it should be emphasized that docking methods are actually tools for "ligand" (rather than "drug") design. The identification of a tight-binding ligand is a necessary but not sufficient criterion toward a promising novel lead structure and its development into a drug. Aspects of synthetic accessibility, bioavailability, or toxicity are not the primary subject of docking, but because it is important to consider these factors at early stages of a design project, filters may be applied to compound libraries before docking or to hits obtained from virtual screening on an early stage. This aspect of pre- or postprocessing is discussed only briefly.

2 GENERAL CONCEPTS AND PHYSICAL BACKGROUND

2.1 Protein–Ligand Interactions and the Physical Basis of Biomolecular Recognition

The selective binding of a small-molecule ligand to a specific protein is determined by structural and energetic factors. For ligands of pharmaceutical interest, protein-ligand binding usually occurs through noncovalent interactions. The physical basis of noncovalent interactions is generally well established through the theories of electromagnetic forces or, on a more fundamental level, of quantum mechanics. For macromolecules, liquid systems, or solutions, however, direct application

of these first principles is significantly complicated by the size and complexity of the systems, in which a large number of fluctuating particles simultaneously interact and influence each other. Principles from classical mechanics and heuristic models are therefore frequently used as an approximation to describe protein–ligand interactions in aqueous solu-

The primary forces acting between a protein and a ligand are all of electrostatic nature. It is the interaction between explicit charges, dipoles, induced dipoles, and higher electric multipoles that leads to phenomena that are commonly referred to as salt bridges, hydrogen bonds, or van der Waals interactions. In simplified classifications, it is only the charge-charge interaction that is called electrostatic. This interaction between two charges is of long range and considerable strength. In vacuum or uniform media it can be described by Coulomb's law. In aqueous solution of biomolecules, however, its application is complicated because of the presence of a large number of water molecules. Unless a sufficiently large number of water molecules is explicitly included in the calculations [as usually only tractable in computationally expensive molecular dynamics simulations (31)], the correct treatment of electrostatic interactions in solution requires solving the Poisson-Boltzmann equation, where the solvent is considered as a continuous medium of high dielectric constant surrounding a low-dielectric solute (32).

Electrostatic interactions, however, do not only occur between charge monopoles. In a comprehensive treatment of electrostatics one has to consider a full power series, and therefore interactions between higher electric moments, such as dipoles and quadrupoles, also play an essential role. Their interaction energies are orientation dependent and become shorter in range with increasing electric moment. For example, in contrast to the $1/r$ dependency in Coulomb's law, the energy of the interaction between a charge and a dipole decays with $1/r^2$, the interaction between two dipoles with $1/r^3$. This, however, is valid only for a fixed orientation of the dipoles. If they are mobile, as in isotropic media (liquids), the dipole-dipole interaction is thermally averaged and an average interaction proportional to $1/r^6$ results. This $1/r^6$ dependency is also encountered in interactions that arise between induced electric moments, such as the dispersion interaction based on London forces. The attractive interactions between (induced) electric multipoles are generally summarized in the term van der Waals interactions. Accordingly, van der Waals forces are weak, attractive, short-range forces that decay with $1/r^6$. These are normally described by intermolecular interaction potentials such as the Lennard-Jones potential:

$$E \propto A/r^{12} - B/r^6$$

where A and B are parameters depending on the type of the interacting atoms. The r^{12} term reflects the short-range repulsive forces attributed to unfavorable spatial overlap of electron clouds at short distances.

An interaction deserving special attention is that of hydrogen bonds (33, 34). In principle, their origin is of the same nature as the interactions mentioned above. A hydrogen bond is defined as the interaction of an electronegative atom (the hydrogen-bond acceptor) with a hydrogen atom covalently bonded to an electronegative atom (the hydrogen-bond donor). The major component of a hydrogen bond is the electrostatic interaction of the donor-hydrogen dipole with the negative partial charge of the acceptor. The special characteristics originate from the fact that the hydrogen atom is very small and can bear a considerable positive partial charge, such that the acceptor can contact the hydrogen atom at a shorter distance than expected from the van der Waals radii. Hydrogen bonds are directed interactions showing a high angular dependency. This directionality arises from the anisotropic charge distribution around the acceptor atom (lone pairs) and the fact that the electron shells of donor and acceptor atom start to overlap at these short distances unless the ideal geometry is maintained. Hydrogen bonds are attributed an important role with respect to specificity of the protein–ligand interaction. This is based on their directionality and the fact that they require a well-defined complementarity in the complex (mutual arrangement of hydrogen-bond donors and ac-

ceptors). However, the importance of hydrogen bonds should not be overemphasized because it is the balance between hydrogen bonds and other forces in protein–ligand complexes that must be appropriately considered (35).

Weakly polar interactions in proteins and protein–ligand complexes are frequently phenomenologically analyzed and classified in terms of the interacting partners (36). This especially includes interactions with π-systems, such as the NH-T, OH-T, or CH-π interaction (37, 38), aromatic-aromatic interactions (parallel π-π stacking versus edge-to-face interaction), and the cation-T interaction (39). All of these can mostly be rationalized in terms of electrostatic interactions outlined above; that is, they involve interactions between monopoles, dipoles, and quadrupoles (permanent and induced). A more distinct character can be attributed to metal complexation, which can play a significant role in individual cases of protein–ligand interactions, as for example in metalloenzymes (2, 40, 41).

Finally, so-called hydrophobic or lipophilic interactions are often mentioned as additional contribution to protein–ligand interactions. These terms are used to describe the preferential association of nonpolar groups in aqueous solution. It should be emphasized, however, that in contrast to what the name suggests, there is no special hydrophobic force. Instead, one should speak of a hydrophobic effect. As further mentioned below, according to the generally accepted view, it arises primarily from the entropically favorable replacement and release of water molecules (42, 43). The association between the nonpolar surfaces itself is simply based on weak London forces (36).

Thermodynamically, the strength of the interaction between a protein and a ligand is described by the binding affinity or (Gibbs) free energy of binding. Assuming a simple equilibrium reaction of the form

$$P + L \rightleftharpoons PL$$

between a protein P and ligand L to give the complex PL, the dissociation constant K_d (or binding constant K_i) is generally used to describe the stability of complex formation:

$$K_d = [P][L]/[PL].$$

From the experimentally measured equilibrium constant the binding affinity can be calculated as

$$\Delta G^0 = RT \ln K_d$$

where R is the gas constant (8.314 J/molK) and T is the temperature [the equilibrium constant would actually have to be related to a standard concentration to become a dimensionless quantity, but in general this is not explicitly considered (44, 45)]. Experimentally determined binding constants K_i (K_d) are typically in the range of 10^{-2} to 10^{-12} M, corresponding to a Gibbs free energy of binding of roughly −10 to −70 kJ/mol (1, 2).

According to the Gibbs-Helmholtz equation, the free energy of binding consists of an enthalpic and an entropic contribution:

$$\Delta G = \Delta H - T\,\Delta S$$

The enthalpy and entropy of binding can be determined experimentally, as, for example, by isothermal titration calorimetry (46,471. These data, however, are still sparse and not always easy to interpret (48, 49). Substantial compensation between enthalpic and entropic contributions is observed (50–52); this phenomenon and its interpretations have recently been critically reexamined (53). Interestingly, the data also show that binding can be both enthalpy-driven (e.g., streptavidin-biotin, AG = −76.5 kJ/mol, AH = −134 kJ/mol) or entropy-driven (e.g., streptavidin-HABA, AG = −22.0 kJ/mol, ΔH = +7.1 kJ/mol) (54). However, because of strong temperature dependencies, even this partitioning is a question of the temperature used for measuring.

What are the major contributions to the enthalpy and entropy of binding? Direct interactions between the protein and the ligand are obviously very important for the enthalpy of binding. Besides that, an essential factor is that protein–ligand interactions occur in

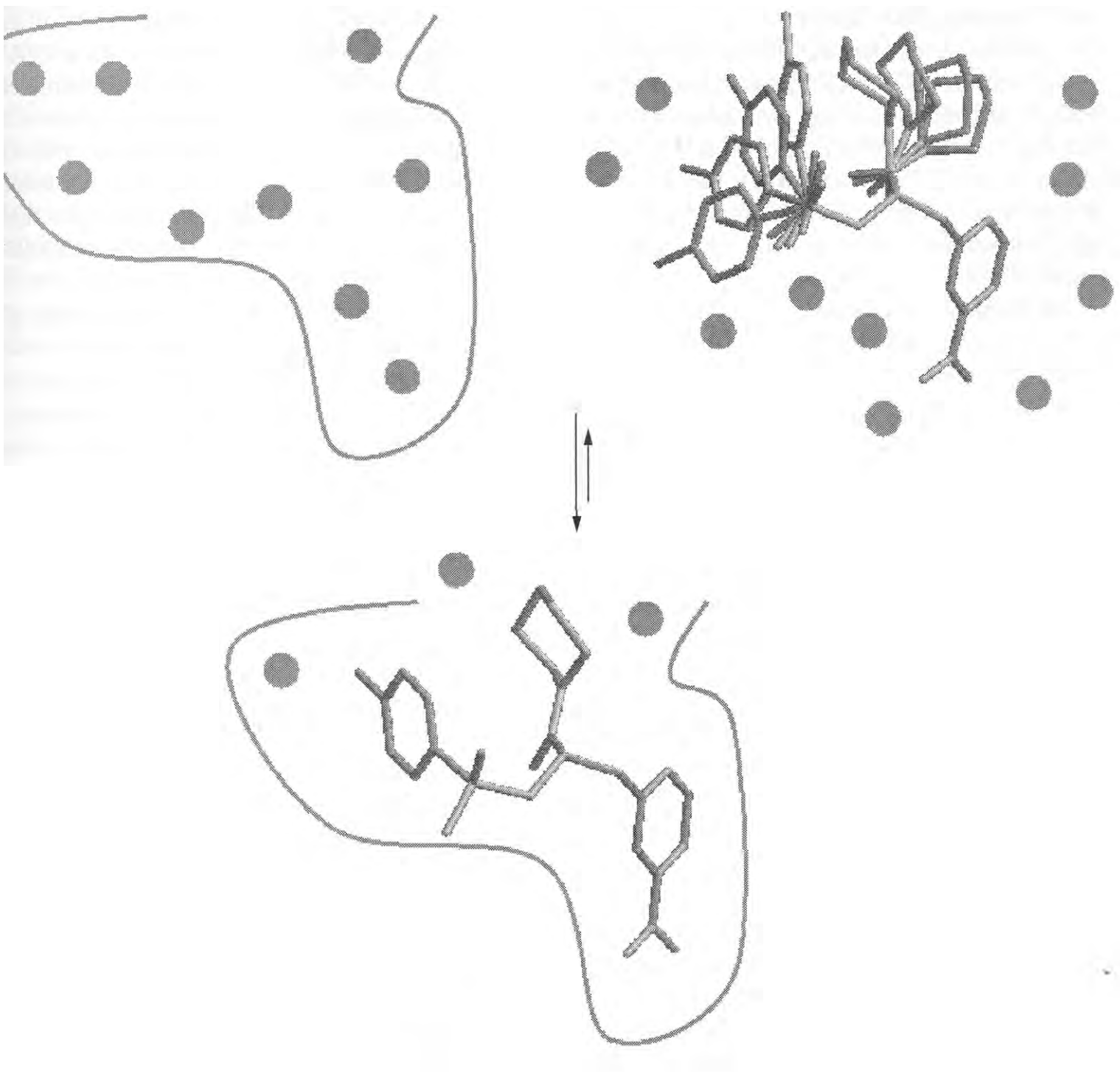


Figure 7.1. Overview of the receptor-ligand binding process. All species involved are solvated by water (symbolized by gray spheres). The binding free energy difference between the bound and unbound state is a sum of enthalpic components (breaking and formation of hydrogen bonds, formation of specific hydrophobic contacts) and entropic components (release of water from hydrophobic surfaces to solvent, loss of conformational mobility of receptor and ligand).

aqueous solution (cf. Fig. 7.1). The unbound reaction partners are solvated and partial desolvation is required before complex formation can occur. This is important for the enthalpy balance because the net energy gain upon complexation can only be the difference between the direct protein–ligand interaction enthalpy and the desolvation enthalpies of the two molecules. In this context, the hydrophobic effect has to be considered again. Upon the formation of lipophilic contacts between apolar parts of the protein and the ligand, unfavorably ordered water molecules are replaced and released. This leads to an entropy gain that is attributed to the fact that the water molecules are no longer positionally confined. In addition, there is an enthalpic contribution: water molecules occupying lipophilic binding sites are unable to form hydrogen bonds with the protein, but after release they can form strong hydrogen bonds with bulk water. Because the removal of hydrophobic surfaces from contact with water leads to negative changes in the heat capacity (ΔC_p), the buried hydrophobic surface area has frequently been correlated with ΔC_p values measured upon li-

gand binding. This, however, may be an oversimplification, neglecting other potential contributions to ΔC_p (55). As further noted by Tame, enthalpy-entropy compensation and the temperature dependency of ΔH and $T\Delta S$ (which are both directly related to ΔC_p), make it ultimately impossible to consider polar or apolar contributions as purely enthalpic or entropic, respectively (56).

Entropically unfavorable contributions arise from the loss of translational and rotational degrees of freedom upon complexation, whereas a small gain in entropy can result from low-frequency concerted vibrations in the complex. A more important factor to consider in an actual design process is conformational flexibility. Upon binding, internal degrees of freedom are frozen, the ligand loses a considerable amount of its flexibility, and usually binds in one single orientation. This is also the explanation why rigid analogs of flexible ligands show higher affinity, as, for example, observed for cyclic derivatives of ligands that adopt the same binding mode as the open-chain derivative (57, 58). Accordingly, higher affinity also results if the protein-bound ligand conformation is already preorganized in solution.

From a variety of experiments, quantitative estimates for some of the mentioned energetic contributions to protein-ligand binding could be derived. Based on data from protein mutants, the contribution of individual hydrogen bonds to the binding affinity has been estimated to be 5 ± 2.5 kJ/mol (59–62). This is similar to what has been obtained for the contribution of an intramolecular hydrogen bond to protein stability (63, 64). The consistency of values derived from different proteins suggests some degree of additivity in the hydrogen-bonding interactions. The accurate description of the interplay with water molecules remains, however, a most challenging task. The contribution of hydrogen bonds to the overall affinity strongly depends on local solvation and desolvation effects and can sometimes be very small or even adverse to binding, as illustrated by the comparison of ligand pairs differing by just one hydrogen bond (65). Charge-assisted hydrogen bonds are stronger than neutral ones, but also associated with a higher desolvation penalty. Thus, the electrostatic interaction of an exposed salt bridge contributes as much as a neutral hydrogen bond (5 ± 1 kJ/mol according to Ref. 66), but the same interaction in the interior of a protein can be significantly stronger (67). Because of the complicated interplay with water, a detailed analysis of the thermodynamics of hydrogen bond formation can sometimes yield surprising results. For a particular hydrogen bond in complexes of-the FK506-binding protein, it has been found that its formation is enthalpically unfavorable but entropically favorable (60). The entropy gain appears to be attributable mainly to the replacement of two water molecules (68).

Contributions from hydrophobic interactions have frequently been found to be proportional to the lipophilic surface area buried from solvent, with values in the range of 80–200 J/(mol Å^2) (69–71). The entropic penalty for freezing a single rotatable bond has been estimated to be 1.6–3.6 kJ/mol at 300 K (72, 73); recent estimates derived from NMR shift titrations are much lower (0.5 kJ/mol) (74), but in the systems studied the conformational restriction may not have been as high as in a protein binding site. Finally, the unfavorable entropy contribution from the loss of translational and orientational degrees of freedom has been estimated to be around 10 kJ/mol (75, 76).

Despite many inconsistencies and difficulties in interpretation, most of the experimental data suggest that simple additive models of protein–ligand interactions might be a reasonable starting point for the development of methods to predict binding affinities, that is, for the derivation of empirical scoring functions. Still, it has to be kept in mind that the assumption of additivity in biochemical phenomena is not strictly valid (77). On the other hand, the large body of experimental data on 3D structures of protein–ligand complexes and binding affinities allows one to derive some general characteristics about protein–ligand interactions. Several features are commonly found in complexes of tightly binding ligands:

1. A high steric complementarity between the protein and the ligand, an observation often described as the lock-and-key paradigm (78,

79). This complementarity, however, is frequently not the result of a match between rigid bodies, but rather achieved through significant conformational changes of both binding partners, a phenomenon generally referred to as induced fit. Additionally, electrostatic complementarity can also be induced, for example, by strong pK_a shifts upon ligand binding that result in the release or uptake of protons of different functional groups either of the protein or the ligand.

2. A high complementarity of the surface properties. Lipophilic parts of the ligands are generally in contact with lipophilic parts of the protein, whereas polar groups are usually paired with suitable polar protein groups to form hydrogen bonds or ionic interactions.
3. An energetically favorable conformation of the bound ligand. Significant conformational strain is usually not observed in ligands binding with high affinity.

In addition to insights taken from high-affinity complexes, experimental information about weakly bound complexes could be equally instructive. Such information has indeed been recognized to be vital for the development of scoring functions (80). Structural data on unfavorable protein–ligand interactions, however, are sparse, partly because structures of weakly binding ligands are more difficult to obtain and are usually considered less interesting by many structural biologists. What can be concluded from the available data is that an imperfect steric fit at the lipophilic part of the protein-ligand interface leads to reduced binding affinity and that unpaired buried polar groups at the protein–ligand interface are strongly adverse to binding. Few buried CO and NH groups in folded proteins fail to form hydrogen bonds (81). Therefore, in the ligand design process an important prerequisite to be regarded is that polar functional groups, either of the protein or the ligand, will find suitable counterparts if they become buried on ligand binding.

.2 Docking, Scoring, and Virtual Screening: e Basic Concepts

he subject of docking is the formation of nonvalent protein–ligand complexes. Given the structures of a ligand and a protein, the task is to predict the structure of the resulting complex. This is the so-called docking problem. Because the native geometry of the complex can generally be assumed to reflect the global minimum of the binding free energy, docking is actually an energy-optimization problem (82), concerned with the search of the lowest free energy binding mode of a ligand within a protein binding site. The macromolecular nature of the protein and the fact that binding occurs in aqueous solution complicate the problem significantly because of the high dimensionality of the configuration space and considerable complexity of the energetics governing the interaction. Accordingly, heuristic approximations are frequently required to render the problem tractable within a reasonable time frame. The development of docking methods is therefore also concerned with making the right assumptions and finding acceptable simplifications that still provide a sufficiently accurate and predictive model for protein–ligand interactions.

Regardless of the nature of the interacting partners, computational docking always requires two components, which may briefly be characterized as "searching" and "scoring" (83). "Searching" refers to the fact that any docking method has to explore the configuration space accessible for the interaction between the two molecules. The goal of this exploration is to find the orientation and conformation of the interacting molecules corresponding to the global minimum of the free energy of binding. Unless the degrees of freedom are restricted to translation and rotation by treating both molecules as rigid bodies, a full systematic search of all "dockings" is normally not feasible because of the huge number of potential solutions and the large amount of computational resources needed to evaluate them. Different strategies are therefore required, which should be accurate and efficient: accurate in the sense that the optimization procedure should not miss any valuable solution (near-global minima), and efficient in terms of computing time and with respect to the fact that the algorithm should not spend unnecessary time by exploring irrelevant regions or by rediscovering previously detected local minima. As will be elaborated in the next

section, there are two opposing approaches to simplify the docking problem either by reformulating it to a discrete problem that can be solved with combinatorial algorithms or by using stochastic search algorithms.

"Scoring" refers to the fact that any docking procedure must evaluate and rank the configurations generated by the search process. The scoring scheme most closely related to experiment, the *ab initio* calculation of the free energy of binding, is not easily accessible to computation. Hence, approximate scoring functions must be used that model the binding free energy with sufficient accuracy and correlate well with experimental binding affinities. In particular, the scoring function should be able to discriminate between native and nonnative binding modes.

Scoring is actually composed of three different aspects relevant to docking and design:

1. Ranking of the configurations generated by the docking search for one ligand interacting with a given protein; this aspect is essential to detect the binding mode best approximating the experimentally observed situation.
2. Ranking different ligands with respect to the binding to one protein, that is, prioritizing ligands according to their affinity; this aspect is essential in virtual screening.
3. Ranking one or different ligands with respect to their binding affinity to different proteins; this aspect is essential for the consideration of selectivity and specificity.

If one were able to accurately calculate the free energy of binding, all three aspects would be satisfied simultaneously. Current scoring functions used in docking programs, however, can usually resolve satisfactorily only the first aspect. They provide only a rough estimate with respect to the comparison across different ligand or protein systems. This is the case whenever the scoring scheme neglects certain factors that are virtually constant for different binding modes with respect to one protein, but that matter for comparisons with other proteins.

Following the general paradigm shift in structure-based design from single compounds to compound libraries, state-of-the-art docking and scoring methods have to be sufficiently fast to be applied for virtual screening. The general strategy of a virtual screening process based on the 3D structure of a target typically involves the following steps:

1. Analysis of the 3D protein structure.
2. Selection of key interactions that need to be satisfied by all candidate molecules.
3. Computational search in chemical databases for compounds that potentially satisfy the key interactions, fit into the binding site, and form additional interactions with the protein; this is done by means of docking and/or structure-based pharmacophore searches.
4. Postprocessing by analyzing the retrieved hits and removing undesirable compounds.
5. Synthesis or ordering of the selected compounds.
6. Biological testing, eventually crystallographic confirmation.

All these steps will be discussed in some more detail in section below. Of primary interest in the context of this chapter is step 3. It requires high-throughput docking with efficient search algorithms, and scoring functions that are able to provide a good separation between potentially "binding" and "nonbinding" ligands. The database or library that is screened should consist of a sufficiently large and diverse set of relevant compounds. Thus, library design is increasingly applied to ensure that only reasonably preselected compounds are docked (29, 84, 85).

3 DOCKING

In this section, approaches to the docking problem are presented with respect to the docking algorithm and the search aspect, Scoring is discussed separately in Section 4. It should be noted in this context, that although a specific docking method is frequently associated with a certain scoring procedure, many docking methods could in principle be combined with a variety of different scoring functions, either for postprocessing of the results

or as objective function during the optimization. Actually, such strategies are followed by considering multiple scoring schemes to achieve "consensus scoring" (86) or "multidimensional scoring" (87). The emphasis in this section is on general characteristics and principles, rather than individual methods, although occasionally specific docking programs have been selected as representative examples for a more detailed illustration of a general concept. The interested reader is referred to Table 7.1 for an overview of currently used docking programs described in the literature. In addition, a valuable source of information is the corpus of regularly published reviews in the field of docking (18, 19, 26, 27, 83, 88–95).

3.1 General Concepts to Address the Docking Problem

Essential for any docking method is a search algorithm that samples the configuration space of two interacting molecules. These molecules need to be represented in a way that is suitable for efficient handling by the search algorithm. Docking methods may therefore roughly be classified by the way the macromolecular receptor is represented (Section 3.1.1), by the handling of the ligand (Section 3.1.2), and—most important—by the search algorithm itself (Section 3.1.3).

3.1.1 Representation of the Macromolecular Receptor. The most straightforward approach for representing the macromolecular structure in a docking application would be by atomic coordinates of the entire protein. A full atomic representation, however, is generally impractical because of the size and complexity of protein structures. The structural information therefore needs to be reduced to a manageable yet representative size and form.

A first step into this direction is to limit the search area to the region surrounding the putative binding site. This is general practice in protein-ligand docking (whereas in protein-protein docking often the entire surfaces are searched for appropriate matches). Scanning of the entire surface for potential binding regions of a small-molecule ligand would hardly be feasible with most docking methods. Furthermore, it would be rather unreasonable to ignore information already available from biochemical experiments or structural data of related complexes. If no such information is available [a situation that we may increasingly be facing as a consequence of the effects of the structural genomics initiatives (17, 96)], methods to identify binding sites are required before the actual docking process can start. Examples are programs for geometric cavity detection, such as LIGSITE (97) or PASS (98), tools to infer protein function from structural homologies (99, 100), or more sophisticated approaches based on a physicochemical and geometrical characterization of binding sites (101). Some docking programs incorporate routines for binding site identification as preprocessing steps (102).

Despite a reduction to only a specified part of the protein surface, a simple representation in terms of atomic coordinates is not practical for most docking procedures. Instead, the space available for ligand binding is frequently characterized by other means that permit more efficient searches. A first alternative is given by geometric shape descriptors, sometimes combined with a physicochemical description. Approaches of this class include molecular surface cubes (103), surface normals at sparse critical points (104), and modified Lee-Richard's dotted surfaces, with each dot coded by chemical property and accessibility (105). A further prominent example is the sphere images of the binding site used in DOCK (106, 107). These spheres are complementary to the molecular surface and represent a space-filling negative image of the binding site. Another important concept that goes beyond a pure geometric description and represents interaction properties of physicochemical relevance is the usage of interaction sites or points, as introduced by the program LUDI (108,109). These interaction sites are discrete positions and vectors in space serving as dummy representations for atoms capable of forming hydrogen bonds or filling hydrophobic pockets. The docking tool FlexX is based on this concept (110). Also, the program SLIDE (111) and the new approach by Diller and Merz (112) use interaction points for fast docking.

A popular alternative to geometric or physicochemical descriptors is the grid representa-

Table 7.1 Overview of Currently Used Programs for Protein–Ligand Docking

Class of Docking Method[a]	Name of Program	Year Published	Original References	Selected References to Further Developments and Applications
Geometric/combinatorial				
Shape/descriptor matching	DOCK	1982	(106)	(11, 127, 143, 371)
	FLOG	1994	(121)	(366)
	ADAM	1994	(135)	(386)
	LIGIN	1996	(387)	(234)
	SANDOCK	1998	(105)	(13)
	QSDOCK	2000	(136)	
	SLIDE	2000	(111)	
	FRED	2001	(123)	
	(Diller & Merz)	2001	(112)	
Incremental construction	FlexX	1996	(110)	(130, 138, 139, 216, 233)
	Hammerhead	1996	(388)	
	DOCK4.0	1998	(328)	(131)
Systematic search (transl. + rot.)	EUDOC	2001	(125)	(389)
Energy-driven/stochastic				
Monte Carlo simulated annealing	AutoDock	1990	(113)	(95, 115, 390)
	RESEARCH	1992	(146)	(145)
	MCDOCK	1999	(147)	
Monte Carlo minimization	ICM	1994	(116)	(82, 117, 201)
	(Caflisch et al.)	1997	(150)	(151)
	QXP	1997	(152)	
	PRODOCK	1998	(119)	(118)
Molecular dynamics (MD)	MDD	1994	(164)	(165)
	(Luty et al.)	1995	(169)	
	(Vieth et al.)	1998	(166)	
	q-jumping MD	2000	(167)	(168)
Genetic algorithm	GOLD	1995	(176)	(177)
	AutoDock3.0	1998	(115)	(208, 228, 391)
	GAMBLER	1999	(86)	
	DARWIN	2000	(178)	
Tabu search	PRO_LEADS	1998	(188)	(189, 360)
Tabu search + genetic algorithm	SFDock	1999	(392)	
Eigenvector following	Low Mode Search	1999	(211)	
Mining Minima algorithm	Mining Minima	2001	(190)	

[a]The classification provided in the first column can only be approximate for programs that offer a variety of different functionalities or follow multistep strategies.

tion of protein structures. The general principle of this approach is that the protein is represented by a set of affinity grids or maps that cover the entire search region. These regularly spaced, orthogonal grids are calculated before the actual docking process. At every grid point, some sort of scoring value or interaction energy of a probe atom with the entire

protein is calculated, providing a map of pseudo-affinities for each atom type or interaction type possibly present in the ligands to be docked. These maps then serve as look-up tables for the calculation of the interaction energy or scoring value during the docking process. Examples of docking programs using this approach are AutoDock (113–115), ICM (82, 116, 117), or ProDock (118, 119).

It should be noted that most of the mentioned representations of protein structure imply that the protein remains rigid during the docking process. As a matter of fact, docking under the assumption of a rigid protein is still common practice in standard applications. Although an acceptable simplification under certain circumstances, it can represent a serious limitation if only unbound protein structures are available. As a consequence, the inclusion of protein flexibility in the docking process is an active area of research, and a separate section is dedicated to this issue (cf. Section 3.2.1).

3.1.2 Ligand Handling. For the ligand, a complete representation in atomic coordinates is perfectly feasible. Ligand atoms may be used directly for matching with binding site descriptors or in the calculation of interaction energies in the case of energy-driven procedures. The central problem is conformational flexibility. Predicting the binding conformation of a ligand is in fact a major component of the docking problem, given that this conformation can significantly differ from that adopted in other environments.

Two general strategies for ligand handling may be distinguished: whole-molecule approaches and fragment-based methods. In the first case, the ligand is docked as an entire molecule. This is rather straightforward if the ligand is treated as a rigid body and only translational and rotational degrees of freedom are considered. Such rigid docking was common practice in early docking algorithms (106, 120). A straightforward extension to account for flexibility is to separately dock precalculated conformers of a given molecule (variant 1 in Fig. 7.2). Explicit docking of multiple conformers has, for example, been obtained with the FLOG program (121). FLOG deals with conformational flexibility by generating different conformers using distance geometry and docking each conformer in a rigid-body fashion. A similar approach has also been obtained with the DOCK program (122). To avoid redundancy in the docking, a common rigid fragment is identified, which is docked only once for the entire set of pregenerated conformers. The flexible portions of the molecule that determine the different conformations are subsequently scored based on the preplacement of the rigid fragment. Yet other examples for rigid docking of multiple conformers are provided by the programs FRED from OpenEye Scientific Software, which performs a fast exhaustive search over all possible orientations (123), and SYSDOC (124) or EUDOC (125), which use fast affine transformation to perform systematic searches over the translational and rotational degrees of freedom of the ligand.

Although this multi-conformer docking can be efficient and accurate for molecules with a limited number of discrete, low-energy conformations, it is less suited for larger and highly flexible molecules, simply because the number of possible conformations increases dramatically. Another way of partially accounting for conformational flexibility in whole-molecule rigid-body docking is to subject the initial matches to some kind of optimization that allows for conformational relaxation. This could be done with some standard energy minimization technique (126,127) or other procedures that resolve clashes of the initial placement by rotation about single bonds, as done, for example, in the docking program SLIDE (111).

A more rigorous treatment of ligand flexibility in whole-molecule docking is performed by sampling ligand conformation space during docking (variant 2 in Fig. 7.2). It normally requires ligand conformational energies to be evaluated besides intermolecular interaction energy. Molecular mechanics force fields are frequently applied for this purpose. Although a more exhaustive sampling of accessible conformations within the binding site is definitely achieved, an obvious disadvantage is the higher computational demand and possibly a reduced efficiency of the algorithm because of lengthy exploration of local minima.

An interesting variant of whole-molecule representations is the use of internal coordi-

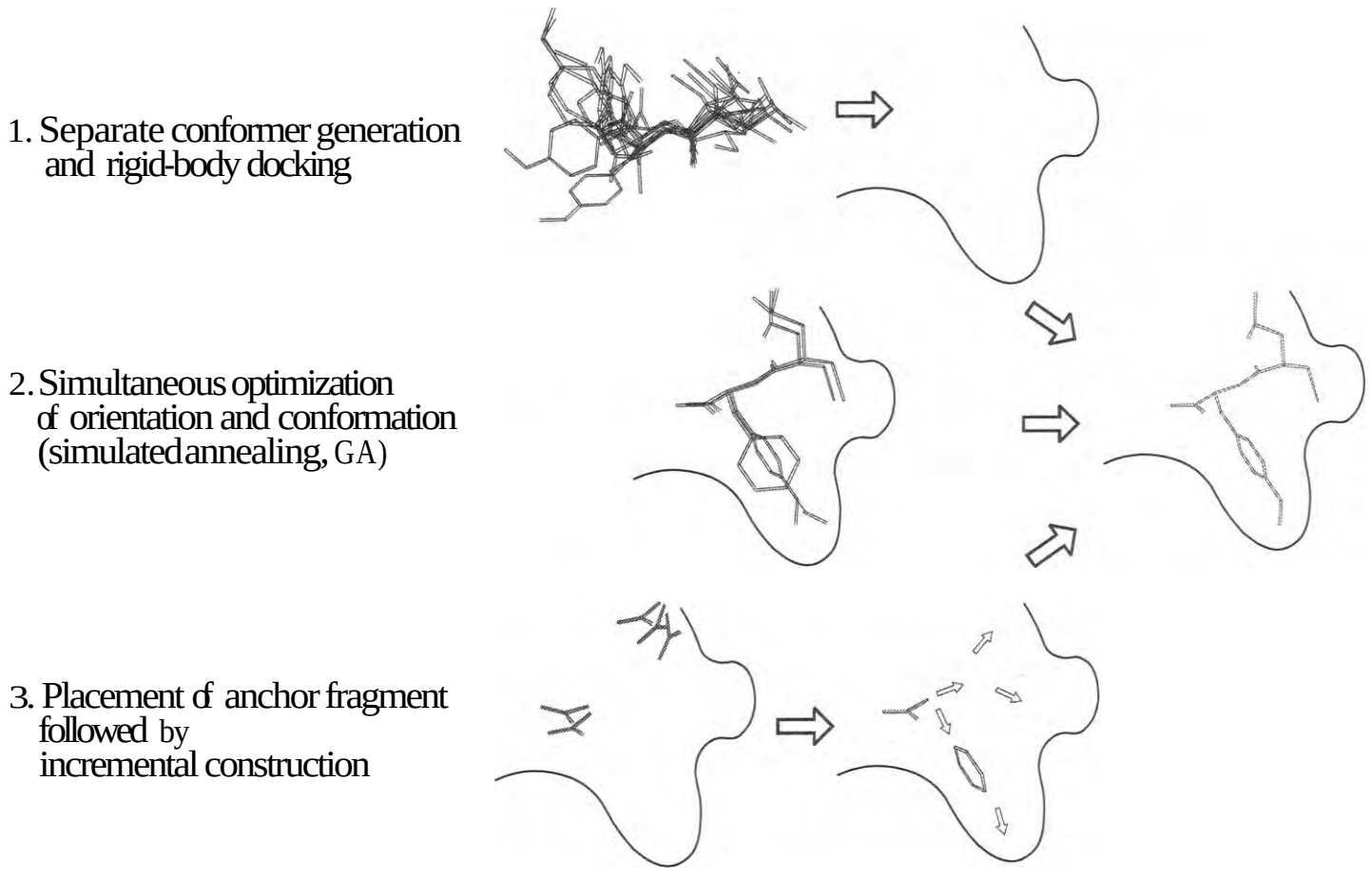


Figure 7.2. Strategies for flexible ligand docking.

nates instead of Cartesian coordinates (82). Internal coordinates help to reduce the number of variables defining the conformation of the molecular system. In Cartesian space, three functionally equivalent variables per atom are required. Internal coordinates, instead, consist of bond lengths, bond angles, and torsion angles. Because bond lengths and angles can be considered rigid to a good approximation, only the torsion angles matter as variables to map conformation space. An efficient implementation of docking algorithms operating on internal coordinates has been obtained, for example, with the ICM method (82, 116, 117).

Fragment-based techniques are an alternative to whole-molecule docking (variant 3 in Fig. 7.2). Here, the molecule is dissected into fragments that can be docked individually in a rigid fashion. The fragments can either be docked separately and then reconnected, or the ligand is built up incrementally following a certain fragmentation scheme. The first variant is very common to programs dedicated to *de novo* design rather than pure docking. Methods of this class have been reviewed extensively (26, 27). However, the approach has also been applied for docking (128) and compared to the whole-molecule docking approach (129).

The other variant of fragment-based ligand docking is used in incremental construction algorithms (110, 130), sometimes also referred to as "anchor and grow" (131). These search strategies are further described below. They dissect the ligand into modular portions and rebuild it incrementally within the binding site starting from the docking position of a suitable base fragment. The advantage is that many potential combinations are eliminated early in the construction, but success critically depends on the selection and placement of the base fragment.

3.1.3 Strategies for Searching the Configuration and Conformation Space. Search strategies of automated docking procedures may roughly be classified as geometric or combinatorial on the one hand and energy driven or stochastic on the other, although ultimately

all methods try to optimize a function that models to some extent the free energy of binding.

3.1.3.1 Geometric/Combinatorial Search Strategies. Most of the early docking methods were entirely based on the concept of shape complementarity. Until today this is the fundamental idea in most protein-protein docking programs. The observation that protein–ligand complexes frequently show a remarkable shape fit of both binding partners has stimulated the conception of surface or descriptor matching as docking search technique. The molecules are represented by geometric and/or physicochemical descriptors and various alignment procedures are applied to match complementary parts of ligand and protein. An example is the original DOCK method, where the ligand is superimposed onto a negative sphere image of the binding pocket, using a distance matching algorithm followed by least-squares fitting (106, 132). Other examples are the least-squares fitting procedure described by Bacon and Moult to achieve matches between complementary surface patterns (133), or the hierarchical search of geometrically compatible triplets of surface normals on the molecules to be docked, as proposed by Wallqvist and Covell (134). The program ADAM performs a complete combinatorial search over all possible matches between hydrogen bond patterns (135). Recently, a new matching algorithm based on so-called quadratic shape descriptors has been described (QSDock); along with the presentation of their method, the authors also provide an extensive discussion of shape-based docking algorithms (136).

Another recent example of descriptor matching is SLIDE, developed as a tool for ligand database screening by docking (111). The binding site is represented by a template of favorable interaction points onto which ligand atoms are matched during the search. Instead of serving as a purely geometric description, these points address four different types of interactions (hydrogen-bond donor, acceptor, donor/acceptor, or hydrophobic interaction center). The search is then performed such that all triangles of appropriate atoms in the ligand are exhaustively mapped onto triangles of template points with compatible geometry and chemistry. This mapping is used to generate initial placements of molecules in the binding site and followed by a series of steps that refine the initial position, resolve collisions, and consider flexibility of both the ligand and the protein side chains (cf. note on hybrid approaches below). Similarly, the rapid docking approach for library prioritization developed by Diller and Merz (112) is based on rigid-body triplet matching of ligand atoms onto precalculated hot spots; subsequently, pruning is performed to remove any positions with significant steric clash, and the remaining matches are subjected to energy minimization.

Pure descriptor matching is efficient for rigid-body docking only. Flexible docking, in fact, is always faced with the additional problem of a combinatorial explosion of possible conformers depending on the number of rotatable bonds. Systematic searches or explicit consideration of each possible conformation would therefore require enormous computing resources. A popular way to address this problem within the class of geometric/combinatorial docking methods is incremental construction (110, 130, 131, 137). The ligand is dissected into fragments and incrementally reconstructed in the binding site starting from a suitably docked base fragment. To avoid dead-end solutions during construction, multiple placements of the base fragment have to be considered. In addition, it can be useful to perform different fragmentations and hence to use different base fragments as starting points, especially for long and highly flexible molecules. The docking itself, that is, the placement of the base fragment and the attachment of remaining portions, is guided by some descriptor matching procedure.

An example of an incremental construction method is the program FlexX (110, 130, 138, 139). Conformational flexibility is considered using a discrete set of preferred torsion angles about acyclic single bonds, together with multiple conformations for ring systems. These torsion angle preferences are taken from a library compiled from torsional fragments extracted from the Cambridge Structural Database (140). The model of molecular interactions is based on similar rules as implemented in LUDI, originating from a composite crystal-

field analysis (141). For each group capable of forming an interaction, a special contact geometry is defined: the group is placed to a center about which an interaction surface is defined, usually as part of a sphere. Two groups form an interaction if the interaction center of one group coincides with the interaction surface of a counter group. To start with the actual docking process, the ligand is fragmented into components by dissecting at all single bonds that are not part of a cycle. Out of these components suitable base fragments are selected. The base fragment is the first portion of the ligand to be placed into the binding site. This is done by superimposing either triples or pairs of interaction centers constructed around the base fragment with triples or pairs of compatible interaction points generated in the binding region. Normally, a large number of initial placements is generated, which is then reduced either by clustering similar solutions or because of clashes with the protein. Next, the incremental construction of the entire ligand is initiated. Starting with the different base placements, the ligand is built up by stepwise linking of the components in compliance with the torsional database. After hooking up additional fragments, new interactions are searched and a scoring function is used to select the best partial solutions, which are expanded in the following step. This is done until the last fragment has been added and placed to result in the complete ligand. The generated ligand positions are finally stored and ranked according to the predicted binding affinity.

An anchor-and-grow algorithm has recently also been incorporated into DOCK (131). Here, after identification of rotatable bonds, the ligand is fragmented into rigid segments, the largest segment is identified as the anchor, and the remaining segments are organized as layers surrounding the anchor. Then the anchor is docked using geometrical matching. Based on the obtained anchor positions, the conformational search is initiated by adding segments from the innermost layer and proceeding outward. This addition is done according to the accessible torsion angle values along the newly added bond. The default is to use two alternative settings for bonds between two sp^2 hybridized atoms, three between two sp^3 atoms and six between sp^2 and sp^3 atoms. The partial constructs are then locally optimized to minimize the sum of intra- and intermolecular energies and pruned back to an approximately constant size of configurations. Pruning is necessary to cope with combinatorial explosion. It is performed on the basis of the score and the orientation, such that both the best scoring and most deviating orientations are retained from each expansion cycle. Finally, after complete reconstruction of the ligand, the pruned set of binding configurations is again subjected to local energy minimization.

This anchor-and-grow implementation in DOCK represents a combination of a geometric and energy-based approach to docking, due to the intermediate steps of energy minimization. As already encountered for SLIDE **(111)**, such multistep or hybrid approaches are commonly found in current docking protocols. DOCK in general is a prototype of such a program, originally based solely on rigid geometric descriptor matching, later enhanced with a variety of additional features. For example, some degree of flexibility has been introduced into the rigid docking procedure by dissecting the ligand into a small set of rigid fragments that are docked separately and then reconnected **(128).** The concept of geometric shape complementarity has been extended to consider physicochemical complementarity by assigning properties to binding-site spheres and allowing them to match only those ligand atoms that are of complementary character, an approach referred to as "sphere coloring" **(142, 143).** Rigid-body minimization has been introduced as refinement after the initial descriptor-matching step **(126)** or in the variant of on-the-fly optimization using force-field energies precomputed on a grid **(127).** In summary, the combination of different approaches and algorithms to overcome the limitations of every single approach has provided us with steadily improving solutions to the docking problem.

3.1.3.2 ***Energy Driven/Stochastic Procedures.*** As mentioned above, docking is essentially an energy optimization problem because the native binding mode of a ligand can in general be expected to correspond to the global minimum of the binding free energy **(82).** Ac-

cordingly, finding this binding mode by docking corresponds to the identification of the global minimum of the free-energy function. Because the actual free energy of binding is not accessible to computation, approximate energy evaluations or scoring functions are used to guide the search. These functions are required to model the free-energy surface in an appropriate way: although the absolute values are not of relevance for the structural aspect of docking, it is essential that the global minimum of a relative free-energy function models accurately enough the position of the global minimum on the real free energy surface. (It is worth mentioning in this context that in purely geometrical or descriptor-based docking procedures, the central assumption is that the degree of surface complementarity or matching between descriptors is proportional to the interaction energy.)

With a suitable energy function available, docking can be performed by global minimization of the energy with respect to the position, orientation, and conformation of the ligand. However, this apparently straightforward approach bears two fundamental problems, inherently related to characteristics of the energy landscape of protein–ligand interactions: the high dimensionality, which precludes a systematic, exhaustive search; and the ruggedness of the surface, reflected by a large number of local minima. Because of this last aspect, standard energy minimization techniques alone are not useful for docking applications because they can guide the search only to the next local minimum. They *are* used, however, in combination with other techniques and play a valuable role at certain stages of the docking process, primarily to refine docked positions and conformations by exploring the local energy landscape in the vicinity of this position.

To address the docking problem, techniques for a more global exploration of the energy landscape are required. A variety of methods is available, frequently used in the context of other modeling applications and optimization problems as well. Three major classes may be distinguished: Monte Carlo techniques, molecular dynamics simulations, and genetic algorithms. Many different variants exist for all of them and frequently, in docking procedures, they are applied in combination with other techniques.

Monte Carlo methods consist of two essential components that are repetitively applied: a random walk of the ligand through the receptor-near space (i.e., the random displacement along translational, rotational, and/or torsional degrees of freedom), and the evaluation of the new configuration based on the Metropolis criterion (144). This criterion decides whether a new position is accepted and hence on the configuration from where the search will proceed. If the energy of the new docked position (E_{new}) is more favorable (lower) than the energy of the previous position (E_{old}), the new position is accepted. If it is less favorable, the probability P for its acceptance is given by

$$P = \exp[-(E_{new} - E_{old})/kT]$$

where k is the Boltzmann constant and T is the effective temperature. To turn this sampling technique into an efficient optimization method applicable to docking, it has to be combined either with a temperature lowering protocol or with some local minimization steps. The former approach is known as Monte Carlo simulated annealing, the latter as Monte Carlo minimization.

In simulated annealing, the effective temperature T is initially set to a high value and gradually lowered, after a predefined number of Monte Carlo steps has been performed at a given temperature. At high temperatures, a broad region of configuration space is sampled: energy barriers can be surmounted because of the high acceptance probability for less favorable placements. As the temperature is lowered, this becomes less probable and the configuration is optimized more locally. Given the stochastic nature of the process, multiple independent runs are required to assess convergence (this equally applies to many of the methods further described below). Examples of docking programs using Monte Carlo simulated annealing as a search strategy are AutoDock (113–115), RESEARCH (145,1461, and MCDOCK (147).

In Monte Carlo minimization, an additional step is inserted after the random walk before Metropolis evaluation. This step is a

local energy minimization, using techniques such as steepest descent or conjugate gradient. Full local minimization after each random-walk step has been reported to improve the efficiency of the procedure (148, 149). A docking procedure that uses global Monte Carlo minimization is the ICM program of Totrov and Abagyan (82, 116, 117). ICM describes both the relative positions of two molecules and their conformations by a uniform set of internal variables and uses precalculated grids of the interaction energies to speed up calculations. Trosset and Scheraga use Monte Carlo minimization in their ProDock program; computational efficiency is enhanced by a grid-based energy evaluation using Bezier splines, which enables one to evaluate gradients and hence to perform minimization on a 3D grid (118, 119). Further Monte Carlo minimization docking procedures have been reported by Caflisch et al. (150,151). Also, the QXP program of McMartin and Bohacek relies on Monte Carlo techniques combined with energy-minimization procedures (152).

Molecular dynamics (MD) simulations represent another technique to sample configuration space (31, 153–157). Based on Newton's equation of motion and principles of statistical thermodynamics, the standard application of this technique is to analyze flexibility and dynamic properties of molecular systems and to calculate free energies in a theoretically rigorous manner (158–163). With respect to protein-ligand docking, MD simulations could in principle be used to simulate the actual binding process, thus providing a "realistic" view of how the docking process proceeds, although this is computationally still out of reach. In fact, standard MD requires massive computational resources, which limits its application to a small number of selected systems. In the context of docking, the problem is that standard MD is slow in exploring global features (crossing of large barriers and exploration of multiple binding sites); accordingly, MD is essentially limited to the simulation and refinement of already bound complexes. Di Nola et al. have addressed this problem in their MDD (MD docking) algorithm (164, 165). This method separates the ligand's center of mass motion from its internal motions. A separate coupling to different thermal baths for both types of motion of the ligand and the receptor is performed. Because the temperature and the time constants of coupling to the baths can be varied arbitrarily, it is possible to increase the kinetic energy of the center of mass of the ligand without increasing the temperature of the internal motions of receptor and ligand. This allows for complete control of the search rate. The technique was applied to the docking of phosphocholine to antibody McPC603, starting from distinct positions well separated from the actual binding site. After appropriate sampling, the average structure of the complex in the binding region was found to closely resemble the crystal structure. Still, the method remains computationally expensive, and thus it is not yet suited for a large-scale application to practical drug design docking problems.

Other docking applications of MD have been reported as well. In a comparison of a CHARMM-based MD docking algorithm with a Monte Carlo and a genetic algorithm, Vieth et al. have observed a comparatively good performance of the MD search for the five analyzed test cases (166). Pak et al. have recently presented a docking approach based on so-called q-jumping MD (167, 168); its basic idea is to apply a smoothed generalized effective potential to enhance conformational sampling by MD. Luty et al. have combined a grid representation for the bulk portion of the receptor with MD simulations of the ligand in the flexible binding site (169). Multiple-copy simultaneous search methods (MCSS) can help to speed up energy-based searches. They use numerous ligand copies that are transparent to each other, but subject to the full force of the protein (170, 171). Finally, short MD simulations are occasionally used at some stage of a docking procedure, primarily with the purpose of local refinement, as for example in the multistep docking strategy of Wang et al., where the last step is an MD-based simulated annealing (129).

The third major class of search methods are genetic algorithms (GAs), which are widely used for docking purposes. GAs are stochastic optimization methods inspired by the concepts of evolution (172–174). The optimization problem is generally formulated in the lan-

guage of genetics. Initially, a random population is generated in which each member corresponds to a potential solution of the problem. A member of the population is represented by its chromosome, in which the variables to be optimized are encoded. This means that each chromosome contains a number of genes, where the genes correspond to the value of a certain variable or set of variables. In the case of docking, the variables for translation and rotation, as well as the torsion angles of the ligand, are encoded in the chromosome. Genetic operators are then applied to the initial population to generate a new population. In general, these operators are "crossover," by which genes from two distinct chromosomes are interchanged to generate two new individuals, and "mutation," by which a given gene is randomly modified. For each newly generated individual the chromosome is decoded (genotype +phenotype) and the fitness of the individual is evaluated. In the context of docking, this fitness is the interaction energy or docking score. Individuals with better scores receive a higher chance for being selected as members of the new population, and thus a higher chance of survival and reproduction into the next generation. Accordingly, the average fitness increases from generation to generation, until, at some point, the process is terminated (by reaching either a fixed number of generations or a constant fitness of the population). The best individual of this final population represents the solution.

Many different variants and implementations of GAs for docking exist, but the general features are always similar. The application of GAs in drug design and docking has been reviewed by Clark et al. (175). A prominent example of a docking program based on a GA is GOLD (176, 177). A special characteristic of GOLD is the direct encoding of hydrogen bonding motifs in the chromosome representation. Upon chromosome decoding, a least-squares fit is used to optimize the overlap of complementary pairs of hydrogen-bonding sites present in the ligand and the receptor. The newest version of AutoDock contains an interesting variant of a GA, a so-called Lamarckian GA (115). This is the combination of a traditional GA with a local search method to perform energy minimization. At each generation, a user-defined fraction of the population is subjected to such a local minimization. This hybrid algorithm was found to be more efficient than a traditional GA, also implemented in AutoDock. A conceptually similar strategy has recently been implemented into the docking program DARWIN (178). Here, a standard GA is combined with a gradient minimization search strategy through an interface to the CHARMM molecular mechanics program (179). Further GA-based docking methods can be found in the literature (86, 180–183).

Another class of evolutionary algorithms that has occasionally found application in the context of docking is known as evolutionary programming (184). Its main difference with respect to GAs is that there is no recombination (crossover) operator, such that evolution is wholly dependent on mutation. Gehlhaar et al. (185) and Westhead et al. (186) have demonstrated the applicability of evolutionary programming to the docking problem, although in a comparative study other algorithms were found to be more effective (186). A new variant called "family competition evolutionary algorithm" has recently been proposed for docking (187).

Besides the three major classes of energy-driven searches (MD, MC, GA), some further heuristic algorithms and search strategies have been developed or adapted for the docking problem. "Tabu search" was found to perform well in comparison with other algorithms (186) and has thus become the main search strategy of the PRO–LEADS docking program (188, 189). Briefly, the tabu search operates on randomly generated positions that are examined on the basis of a tabu list. This list contains a number of previously generated solutions and serves to impose restrictions on the search process: a random move of the ligand is considered "tabu" if it generates a solution that is not sufficiently different from the stored solutions, unless its energy is more favorable than the energy of the best solution so far. Using these restrictions, the search is prevented from revisiting regions of the search space and the exploration of new areas is encouraged. Ideas from tabu search are also used in the recently described adaptation of the Mining Minima algorithm for protein–

ligand docking (190). Here, an exclusion zone is placed around each energy minimum as it is discovered, to avoid rediscovering it in future docking iterations. Mining Minima itself is based on a variety of optimization techniques to gradually focus a large region of random search to areas around the lowest energy minima.

3.2 Special Aspects of Docking

Besides the general characteristics outlined above, there are a number of special issues associated with the docking methodology that deserve explicit consideration: protein flexibility, water molecules, and objective assessment. In addition, the interplay of docking with QSAR methods and homology modeling is of further interest to highlight the possibilities opened by combined application of standard methods in structure-based drug design.

3.2.1 Protein Flexibility. Proteins are inherently dynamic systems (153,191). A single, fixed conformation, even the average provided by a crystal structure, may not be an adequate representation of the protein, unless the system is very rigid (192). Instead, even under standard equilibrium conditions, the native folded state of a protein is best characterized by a collection or ensemble of energetically nearly equivalent conformations. If the conditions are changed, the local minima and the population of these states may shift, eventually resulting in an observable change of the average structure. Also, the introduction of a ligand corresponds to a change of the environment that may lead to similar effects. Accordingly, the binding conformation of the receptor may already be present in the ensemble of protein conformations (193, 194) and the ligand does not actively deform a fixed state of the protein, as generally inferred from the "induced fit" model.

Whatever the actual mechanism might be, the comparison of experimental protein structures in the ligand-free and in the complexed state frequently shows protein conformational changes induced by or associated with ligand binding (195). The spectrum of phenomena ranges from side-chain rotations to loop rearrangements and the movement of entire domains. Accordingly, in the context of docking it is frequently not justified to neglect protein flexibility (35). If no alternative for docking into the rigid protein is available, at least a protein conformation (possibly from a complex structure) should be used that is compatible with suitable binding modes. Obviously, a preferable docking tool would consider full protein flexibility, but appropriate realization of this goal remains a challenge because of the high dimensionality of protein conformation space. Consideration of protein flexibility also complicates the problem of scoring and selecting the best ligand placement, given the difficulty in accurately evaluating protein conformational free energies in addition to ligand-binding free energies.

Current approaches to the problem of flexible protein docking have recently been reviewed by Carlson and McCammon (196), and more briefly by Abagyan and Totrov (18) and Claussen et al. (197). The methods differ by the degree of flexibility they can cover. The least complex methods are those that model small adjustments of contact residues and side chains in an implicit way using soft docking. The protein itself remains fixed, but either through an adapted geometric representation or using a tolerant scoring function a certain amount of overlap between the protein and the ligand is allowed, emulating some "plasticity" of the receptor. The docking program by Jiang and Kim based on the matching of molecular surface cubes is explicitly based on this soft docking idea (103). Other more recent docking approaches have implemented a soft scoring function (198). The advantage of these simple approaches is that they do not increase the demands on computing time.

The next level is represented by methods that allow for explicit side-chain flexibility. GOLD's genetic algorithm can handle the rotation of a few terminal hydrogen-bond donor and acceptor groups to optimize the hydrogen-bonding network (176, 177). A technique to handle larger side-chain movements is the use of side-chain rotamer libraries, as first demonstrated by Leach. In this approach, heuristic algorithms such as dead-end elimination are used to search the large combinatorial space (199). Schaffer and Verkhivker instead use a rotamer library to first generate likely side-chain conformations, which are then sub-

jected to energy minimization together with the docked ligand (200). Another approach making use of minimization has been described by Apostolakis et al.: after "seeding" the receptor with randomly generated ligand positions that may overlap with the protein, the complex is subjected to minimization, during which nonbonded interactions are gradually switched on, to gently relieve steric overlap by minor conformational changes of the ligand and receptor. The best-ranked solutions are then subjected to further refinement by Monte Carlo minimization (151). Furthermore, the Monte Carlo minimization technique in internal coordinates of the ICM program can sample and optimize side-chain torsions during ligand docking (117, 201). Finally, the docking tool SLIDE allows for some side-chain flexibility at the optimization stage of initial placements. In SLIDE, collisions are resolved by rotations about single bonds in the ligand and the protein side-chains to reduce a maximal number of collisions by minimal conformational changes of both binding partners (111,202).

An alternative to account, in principle, for an arbitrary degree of protein flexibility is the use of protein structure ensembles. The ensembles could be assembled from multiple crystal structures of a given protein, from NMR structure determination, or from trajectories of molecular dynamics simulations. In addition, a rotamer library can be used to create a minimal set of new conformations (203). Whatever the origin of the individual members of the ensemble, each represents a distinct conformational state of the protein, and may eventually correspond to the preferred ligand-binding state. Three different ways to use protein ensembles for docking can be distinguished: in its most straightforward form, docking is carried out sequentially with each member of the ensemble using rigid-receptor docking (124,204–206). Another way is to use a weighted-average representation of the ensemble. Knegtel et al. followed this approach by generating composite grids that were used for scoring within the DOCK program (207). Recently, it has also been tested with AutoDock (208). Broughton has developed another method by combining statistical analysis of a conformational ensemble from short MD simulations with grid-based docking protocols (209). The third and most sophisticated approach to handle protein ensembles is implemented into FlexE, a variant of the FlexX program (197). FlexE is based on a united protein description generated from the superimposed structures of the ensemble. For the parts that differ among the protein structures, discrete alternative conformations are explicitly taken into account on the fly during the incremental construction of the ligand in the binding site. As an important feature, these geometric alternatives are optimally joined to create new valid protein structures in a combinatorial fashion. Thus, conformations of the protein are not limited to those explicitly present in the ensemble, nor are the interactions blurred by averaging over distinct alternative instances, which may correspond to unrealistic protein conformations.

The so-called Low Mode Search (LMOD), originally established as a method for conformational analysis (210), has recently been demonstrated to be applicable also to the problem of docking flexible ligands into flexible protein binding sites (211). To explore the potential energy surface of molecules, LMOD is based on eigenvector following, where eigenvectors correspond to the (low-frequency) "normal modes" of vibration. For the purpose of docking, LMOD has been combined with a limited torsional Monte Carlo movement, as well as random translation and rotation of the ligand.

Generally, however, full consideration of flexibility, either of the binding site or the entire protein, remains the domain of MD simulations. The disadvantage is their high computational demand required to achieve significant sampling. Simplified MD restricted to the binding site has been used by Luty et al., where the bulk of the protein receptor is represented as a grid, whereas a full atomic description is used only for the proximity of the binding site to include flexibility in the docking process (169). The approach of Mangoni et al. mentioned above provides a method for enhanced sampling. It has been used to dock a ligand into a receptor that is treated fully flexible and solvated with explicit water molecules (165). Alternatively, shorter MD runs may be used at intermediate or final stages of a dock-

ing procedure to refine complexes generated by rigid-body docking methods. In this case, however, flexibility is not considered simultaneously to the docking process. It thus only refines solutions from rigid receptor docking and does not enhance the scope of the search for possible binding modes.

3.2.2 Water Molecules. Water plays a crucial role in molecular interactions (212, 213). At the interface of a protein–ligand complex, water molecules can have a significant impact on complex formation, either by mediating or improving specificity and affinity of the interaction. They promote adaptability, thus allowing for promiscuous binding (214). Individual conserved ("structural") water molecules can therefore be crucial for the successful design of new inhibitors. A prominent example is the structural water molecule observed in nearly all HIV protease complexes with substrate-like inhibitors. Attempts at replacing it have guided the design of new tight-binding inhibitors [e.g., (215)]. Instead of the usual implicit modeling of solvation effects, explicit consideration of structural water molecules and water-mediated interactions would therefore be a highly desirable feature in docking methods. Ideally, simultaneously to the ligand placement the docking program should be able to predict whether at a particular site water molecules mediating protein–ligand interactions may preferably reside or whether the displacement of these water molecules by appropriate ligand functional groups would be more favorable. No docking tool is yet available to accomplish this task. Obviously, not only the placement of water molecules is demanding, but especially their energy scoring, resulting from the complicated thermodynamics associated with water interactions.

In principle, MD simulations provide the most natural route to the explicit consideration of water molecules. In the MD docking approach described by Mangoni et al., explicit water molecules are indeed used (165). It was found, though, that the presence of explicit water molecules shields the interactions between the ligand and the receptor. Consequently, different weights were applied to the ligand–receptor and ligand–solvent interactions, respectively, to cope with this complication. Because of the high computational costs, the approach seems affordable only in special cases where the presence of explicit solvent appears important.

An approach to explicitly place water molecules during fast docking has been introduced into FlexX (216). In a preprocessing phase, possible favorable water sites in the binding pocket are calculated and stored. During the incremental construction phase of FlexX, water molecules are switched on at these sites if they provide additional hydrogen bonds to the ligand. Steric constraints produced by these water molecules and the quality of the achieved hydrogen bond geometry are then used to optimize the ligand orientation during the construction process. In several cases, water molecules between protein and ligand could be correctly predicted; however, the overall improvement on the FlexX docking results for a test set of 200 complexes was nearly negligible.

The program SLIDE can consider tightly bound waters while docking potential ligands (111). To select which water molecules to retain and which to remove from the binding pocket before docking, the knowledge-based approach Consolv (217) is applied to determine those waters that are likely to be conserved upon ligand binding and to adjust a, penalty for their displacement. Once these waters have been selected to be initially retained upon docking, SLIDE either translates or discards a water molecule to remove overlap with ligand atoms after the ligand has been docked to the binding site. Displacement of a water molecule is performed only if collisions cannot be resolved by iterative translations. Any displacements by nonpolar ligand atoms are penalized upon scoring. In database screening runs on three different target proteins, this procedure was found to produce reasonable results with respect to water-mediated interactions, but no systematic test has been reported so far.

As long as a simultaneous docking of water molecules and ligands is an unsolved problem, it remains common practice to consider essential water molecules as a fixed part of the binding site. Preplaced water molecules may either correspond to recurrently observed waters found in multiple crystal structures of the tar-

get protein, or to predicted positions based on estimated water affinity potentials suggested by programs such as GRID (218–220). The latter strategy has been applied by Minke et al. using AutoDock (221), showing that successful docking of carbohydrate derivatives to the heat-labile enterotoxin critically depends on the inclusion of water molecules. Examples for the consideration of experimentally observed water molecules as part of the target during docking are the studies of Rao et al. (docking to factor Xa using AutoDock) (222) and Pospisil (docking to thymidine kinase using AutoDock and FlexX) (223). The influence of explicit water molecules in docking was also investigated in the validation study of the new program DARWIN (178). Inclusion of explicit water molecules was essential in some cases, unless interaction energies were calculated with a Poisson-Boltzmann-based implicit solvent model. Yet another example is a search for metallo-β-lactamase inhibitors (14) with the docking program FLOG. Docking was performed with three different configurations of bound water in the active site. The top-scoring compounds showed an enrichment in biphenyl tetrazoles. A crystal structure of one tetrazole not only confirmed the predicted binding mode but also displayed the water configuration that had, retrospectively, been the most predictive one of the three models. Further examples from virtual screening studies are available that show that the inclusion of conserved water molecules in the docking process can dramatically improve the hit rate (15,161.

3.2.3 Assessment of Docking Methods. Docking methods are usually assessed by their ability to reproduce the binding mode of experimentally resolved protein–ligand complexes: the ligand is removed from the complex, a search area is defined around the actual binding site, the ligand is redocked into the protein, and the achieved binding mode is compared with the experimental position, usually in terms of a root-mean-square deviation (rmsd). If the rmsd is below 2 Å, it is generally considered a successful prediction. The vious goal is that such a "near-native" solution is ranked best among the set of ligand poses generated. Virtually any introduction of a new docking method has been accompanied by such a test. The number of complexes used has varied as much as the reported success rates, which are between 10% (224) and 100% (152). Clearly, success rates of 100% are rather a consequence of the limited test set size than a reflection of the mere quality of the docking method.

Numerous critical issues have to be addressed in this context. Validations carried out on very few complexes (≤20) do not adequately assess the scope of the method, particularly if no attempt was made to select a representative set of structures that appropriately covers a broad range of binding features important to protein–ligand complexes. Up to now, only a few docking methods have been assessed on a broad range of complexes [e.g., FlexX (200 complexes) (139), ScoreDock and DOCK (200 complexes) (225), EUDOC (154 complexes) (125), DOCK, FlexX, and DrugScore (100–150 complexes) (226), GOLD (100 complexes) (177), the method of Diller and Merz (using the GOLD test set) (112), and PRO–LEADS (70 complexes) (189)]. In the case of GOLD it has been explicitly mentioned that the test set was selected by a researcher not involved in the development of the algorithm (177). The definition of an objective and relevant reference test set that could serve as standard benchmark for every new docking method would be highly desirable for both user and developer (18). First efforts in developing a database that could be of use in this context have been reported (227). Suitable test sets should cover a sufficient number of highly diverse protein–ligand complexes, including cases that provide some challenge to docking methods (e.g., water-mediated interactions, interactions with metal ions). To test performance with respect to potential induced fit, the structure of the unligated protein or alternative complexes with different bound ligands should be available as well. The test set should comprise fully resolved crystal structures with a resolution of ≤2.5 Å. Complexes with ligands significantly involved in crystal packing contacts should be avoided. Such cases will likely fail in reproducing the experimental binding mode because of missing contacts present only in the packing environment (228). Finally, the importance to study low-affinity or "non-binding" ligands must be ad-

dressed; accordingly, experimental information about the binding geometry and affinity of some weak-binding ligands should also be available.

In addition to the tests usually reported by the authors of a program, comparative studies have been reported on the assessment of different docking and scoring approaches. In part they also address some of the aspects raised above. Westhead et al. have presented a comparison of four heuristic search algorithms (simulated annealing, genetic algorithm, evolutionary programming, and tabu search) (186). In an attempt to provide an unbiased comparison, all algorithms were implemented into the PRO_LEADS program and a single scoring function was used. Other recent examples are the studies of Ha et al., who compared DOCK (using two different scoring functions) and FlexX (229), and, in the context of virtual screening, the work of Bissantz et al., who compared DOCK, FlexX, and GOLD together with seven different scoring functions (230) (cf. also Section 5.2 below).

An unbiased test scenario is guaranteed if researchers are provided with a set of protein–ligand complexes of experimentally resolved, but yet unpublished structure. Two such blind trial competitions have been carried out so far (231, 232). A series of interesting issues regarding docking tests and problems with true predictions have been amply discussed by Dixon (231) and participants in the CASP2 docking competition (117, 145, 233, 234). Unfortunately, the number of targets subjected to such blind tests has so far been rather scarce. A major limitation to such blind comparisons is the availability of experimental data before publication.

3.2.4 Docking and QSAR. As long as the problem of accurate binding free energy prediction on the basis of a given complex geometry has not been resolved (cf. section on scoring functions), computational methods establishing quantitative structure-activity relationships (QSARs) to estimate relative binding affinity differences within a set of ligands remain a pragmatic alternative. Both classical and 3D QSAR methods have been developed as ligand-based approaches (235–237). They rely exclusively on ligand information and try to correlate experimental binding data with features described by a set of relevant descriptors. In 3D QSAR, such as CoMFA (Comparative Molecular Field Analysis), these descriptors are essentially virtual interaction energies (van der Waals and coulombic), calculated using an appropriate probe atom placed at the intersections of a regularly spaced grid surrounding the molecules. The model derived from differences in the various interaction fields provides a quantitative spatial description of those molecular properties that matter for binding. They can be interpreted as a surrogate representation of the binding site. Essential for the success of all 3D QSAR approaches is an appropriate alignment of the ligands: their relative spatial superposition must reflect the differences in binding geometry also experienced at the binding site of the structurally unknown protein. Various strategies have been developed to achieve this goal (235, 236). Increasingly, however, these methods are also applied if the receptor structure is known. This results in "receptor-based 3D QSAR," a combination of a ligand-based QSAR approach with information extracted from receptor structures (238). This additional information is used to generate a ligand alignment based on the experimental or predicted binding, mode of the ligands in the binding site. The standard 3D QSAR techniques are subsequently used to derive a correlation model and to ultimately predict the binding affinity of new, appropriately aligned ligands (239). As a practical advantage, receptor-based 3D QSAR provides important information as to which of the protein–ligand interactions are responsible for the variance in biological activity among the given set of ligands.

Obviously, in the case of known receptor structure, the ligand alignment can be obtained by docking. This strategy has indeed been followed in a variety of studies: it has been used to set up CoMFA models [e.g., (240)] or extended to the Comparative Binding Energy (COMBINE) analysis (241–244), that explicitly exploits receptor information to generate the QSAR descriptors. Furthermore, in a GRID/GOLPE (245) analysis, the model generated with the docking alignment has been compared to the traditional CoMFA model

based on ligand alignment (238, 246); the alignment generated by docking could be shown to exhibit higher relevance.

Another concept to combine docking with QSAR has recently been proposed by Vieth and Cummins in their DoMCoSAR approach (247). DoMCoSAR is used to statistically determine the docking mode that is consistent with a structure-activity relationship, based on the explicit assumption that all molecules exhibit the same binding mode. In a first step, all molecules of a chemical series with common substructure are docked in an unbiased way to the protein binding site and the results are clustered to establish the most favorable docking modes for the common substructure. Subsequently, constrained docking is performed by forcing all molecules to align with the common substructure in the major docking modes. In a final stage, interaction-energy-based descriptors are calculated for all major docking modes. QSAR models are then derived to determine the statistically significant and most predictive set of descriptors and thus the docking mode that is most consistent with a given structure-activity relationship. As noted by the authors, the appeal of this method is that an objective statistical justification for the selection of a binding mode is obtained. This may especially prove useful in cases where the primary docking scores yield nearly degenerate multiple binding modes and a selection of the most representative result is difficult. However, because one alignment is rendered prominent among others for the sake of best agreement with the derived QSAR model, the danger exists that unconsidered or ill-defined descriptors in the QSAR could possibly distort the final or accepted alignment.

3.2.5 Docking and Homology Modeling. In the absence of an experimental protein structure, a homology model may be used for docking and structure-based design. Such a model can be generated by comparative modeling based on homologous proteins of known structure. Obviously, it is most reliable in the regions of highest homology between the templates and the target protein. Although an overall skeleton of the target protein can frequently be obtained with sufficient accuracy, the structural details of the binding site are often beyond the scope of the method. In fact, members of a homologous protein family may show considerable differences in the binding region. Accordingly, homology models may not be sufficiently accurate to apply standard docking tools, and special methods addressing the docking of ligands to low-resolution structures have been presented (248).

Clearly, flexible-receptor docking could help to alleviate the problem. A frequently followed alternative is to refine the initial complex between the protein model and the ligand, most commonly by relaxation with MD simulations (249–251). This may also be combined with free energy calculations to determine the binding mode most consistent with experimental affinity data (252). However, refinement does not overcome the problem that the initial conformation of the model may preclude the binding of certain ligands. This has, for example, been demonstrated by Schapira et al. in a virtual screening for retinoic acid receptor (RAR) antagonists based on an RAR homology model (201). The automatic selection procedure based on flexible ligand docking was followed by optimization of the selected candidates with flexible protein side chains using the ICM program (82,116,117). Nevertheless, some known ligands were repeatedly missed by the screening algorithm because of incompatible binding site conformations. Consideration of side-chain flexibility already in the initial docking simulation was required to accommodate these ligands.

An approach developed especially for the purpose of docking ligands into approximate protein models generated by homology modeling is the DragHome method (253). The binding site is analyzed in terms of putative ligand interaction sites and translated using Gaussian functions into a functional binding-site description represented by physicochemical properties. Similarly, ligands are translated into a description based on Gaussian functions and the docking is computed by optimizing the overlap between the two functional descriptions. The use of "soft" Gaussian functions to describe protein–ligand interactions is one possibility to take into account the limited accuracy of modeled structures for the purpose of docking. The method for generating and op-

timizing ligand orientations relative to the binding-site representation was adapted from the ligand alignment program SEAL (254–256). For a set of different ligands, the generated solutions are analyzed with respect to the mutual ligand alignment. This alignment is then used to generate 3D QSAR models, which in turn can be interpreted with respect to the surrounding protein model. This can highlight inconsistencies and deficiencies present in the model, and thus information which in future developments of the methods is planned to be fed back into a subsequent modeling step to improve the protein model. The idea behind this is that the cycle of docking and alignment, ligand data analysis (3D QSAR), and protein structure modeling should be repeated until self-consistency is achieved. This would provide a protein homology model optimized with respect to the binding site and suitable to obtain consistent docking results.

4 SCORING FUNCTIONS

This section is dedicated to the scoring aspect of the docking problem. Various approaches are discussed that try to capture the essential elements of protein–ligand interactions in computationally efficient scoring functions. The discussion focuses on general approaches rather than individual functions. The reader is referred to Table 7.2 for original references to the most important scoring functions.

4.1 Description of Scoring Functions for Protein–Ligand Interactions

Reversible protein-ligand binding is an equilibrium between the bound state and the unbound state of the binding partners. The rigorous theoretical description requires full consideration of all species involved: the separate solvated protein, the separate solvated ligand, and the solvated complex, in which the binding partners are partially desolvated and form interactions with each other. The quantity of interest to characterize this equilibrium is the free energy of binding. Its most accurate calculations are based on the evaluation of ensemble averages according to principles of statistical mechanics (45). To obtain reasonably accurate values of binding free energies, extensive Monte Carlo or MD simulations are necessary, which require large computational resources. Clearly, this is impractical for standard docking applications. Furthermore, even the most advanced techniques are reliable only for calculating binding free energy differences between closely related ligands (162, 163, 257, 258). However, some less rigorous, but faster and, as experience shows, often not less accurate methods have been developed, that are suitable to handle larger numbers of ligands. For example, continuum solvation models are used to replace explicit solvent molecules at least in the final energy evaluation of the simulation trajectory (259), or linear response theory is applied (260–262), sometimes augmented by a surface term (263).

Scoring functions that can be evaluated fast enough to be applied in docking and virtual screening can only estimate the free energy of binding. They usually take into account only one possible configuration of the receptor-ligand complex and disregard ensemble averaging and explicit properties of the unbound states of the binding partners. Furthermore, all methods share the assumption that the free energy can be decomposed into a sum of terms (additivity). In a strict physical sense, this is not allowed, given that the free energy of binding is a state function, although its components are not (77,264). In addition, simple additive models cannot describe subtle cooperativity effects (265). Nevertheless, it is often useful to interpret receptor-ligand binding in an additive fashion (266–268), and estimates of binding free energy based on the additivity assumption are often accessible at very low computational cost.

Three main classes of fast scoring functions can be distinguished: force field–based methods, empirical scoring functions, and knowledge-based methods. The following sections are dedicated to a separate discussion of each method.

4.1.1 Force Field–Based Methods. An obvious idea to circumvent parameterization efforts for scoring is to use nonbonded energies of existing, well-established molecular mechanics force fields for the estimation of bind-

Table 7.2 Overview of Currently Used Scoring Functions

Type of Function	Name of Function	Year Published	Original References	Selected References to Applications
Force field	Charmm	1998	(274)	
Force field + desolvation	(Schapira, Abagyan et al.)	1999	(280)	
	AMBER + desolvation	1999	(276)	
	Charmm + PB	1999	(393)	
	AMBER + desolvation	1999	(278)	
	MM PB/SA	1999	(343)	(344, 346)
Linear response	LIE	1994	(260)	(261,263, 394)
Simplified free-energy perturbation	OWFEG Grid	2001	(284)	(395)
Empirical	(Wade, Goodford et al.)	1989,1993	(220,298)	GRID (218)
	SCORE1	1994	(294)	LUDI (108, 109); (300,396)
	(Miller, Sheridan et al.)	1994	(121)	FLOG (121)
	GOLD score	1995	(176,177)	GOLD (176,177)
	PLP	1995,2000	(185, 367)	
	FlexX score	1996	(110)	FlexX (110,130, 138, 397)
	VALIDATE	1996	(307)	
	(Jain)	1996	(297)	Hammerhead (388)
	ChemScore	1997	(80)	(295)
	SCORE2	1998	(296)	
	(Takamatsu, Itai)	1998	(398)	
	SCORE	1998	(293)	(225)
	AutoDock3	1998	(115)	(391)
	Fresno	1999	(399)	(230)
	ScreenScore	2001	(287)	
Desolvation terms	HINT	1991	(400)	(308)
	(Zhang, DeLisi et al.)	1997	(305)	
Knowledge-based	SMoG	1996	(313)	SMoG (314)
	BLEEP	1999	(315,316)	(340)
	PMF	1999	(317)	(299, 319, 320, 339,369)
	DrugScore	2000	(226)	(15,318)

ing affinity. In doing so, one substitutes estimates of the free energy of binding in solution by an estimate of the gas phase enthalpy of binding. Even this crude approximation can lead to satisfying results. A good correlation was obtained between nonbonded interaction energies calculated with a modified MM2 force field and IC_{50} values of 33 HIV-1 protease inhibitors (269). Similar results were reported n a study of 32 thrombin-inhibitor complexes with the CHARMM force field (270). In both studies, however, experimental data represented rather narrow activity ranges and covered little structural variation.

The AMBER (271, 272) and CHARMM (179) nonbonded terms are used as scoring function in several docking programs. As mentioned above (Section 3.1.1), protein terms are usually precalculated on a rectangular grid to speed up the energy calculation compared to traditional atom-by-atom evaluations (273). Distance-dependent dielectric constants are

usually employed to approximate the long-range shielding of electrostatic interactions by water (274). However, compounds with high formal charges still obtain unreasonably high scores as a result of overestimated ionic interactions. For this reason, a common practice in virtual screening is to separate databases of compounds into subgroups according to their total charges and rank these groups separately. When electrostatic interactions are complemented by a solvation term calculated by the Poisson-Boltzmann equation (32) or faster continuum solvation models (e.g., Ref. 275), effects of high formal charges are usually leveled out. In a validation study on three protein targets, Shoichet and coworkers observed significantly improved ranking of known inhibitors upon correction for ligand solvation (276). The current version of the docking program DOCK calculates solvation corrections based on the generalized Born (277) solvation model (278). The method has been tested in a study where several peptide libraries were docked into various serine protease active sites (279).

In the context of scoring, the van der Waals term of force fields is mainly responsible for penalizing docking solutions with respect to overlap between receptor and ligand atoms. It is often omitted when only the binding of experimentally determined complex structures is analyzed (280–282).

Very recently, a new contribution to the list of force-field-based scoring methods has been developed by Charifson and Pearlman. This so-called OWFEG (one window free energy grid) method (283) is an approximation to the expensive first-principles method of free energy perturbation (FEP). For the purpose of scoring, an MD simulation is carried out with the ligand-free, solvated receptor site. During the simulation, the energetic effects of probe atoms on a regular grid are collected and averaged. Three simulations are run with three different probes: a neutral methyl-type atom, a negatively charged atom, and a positively charged atom. The resulting three grids contain information on the score contributions of neutral, positively, and negatively charged probe atoms located in various positions of the receptor site. They are used for scoring a ligand position by linear interpolation based on the partial charges of the ligand atoms. This approach seems to be successful for K_i prediction as well as virtual screening applications (284). Its conceptual advantage is the implicit consideration of entropic and solvent effects and some protein flexibility.

The calculation of ligand strain energy traditionally also lies in the realm of molecular mechanics force fields. Although effects of strain energy have rarely been determined experimentally (3), it is generally accepted that high-affinity ligands bind in low-energy conformations (285, 286). If a compound must adopt a strained conformation to fit into a receptor pocket, this should lead to a less negative binding free energy. Strain energy can be estimated by calculating the difference between the global energy minimum of the unbound ligand and the current conformation of the ligand in the complex. However, force field estimates of energy differences between individual conformations are not reliable for all systems. In practice, better correlation with experimental binding data is observed when strain energy is used as a filter to weed out unlikely binding geometries rather than including it in the final score. Estimation of ligand strain energy based on force fields can be time-consuming and therefore alternatives are often employed, such as empirical rules derived from small-molecule crystal data (140). Conformations generated by such programs are, however, often not strain-free because only one torsional angle is regarded at a time. Some strained conformations can be excluded when two consecutive dihedral angles are taken into account simultaneously (287).

4.1.2 Empirical Scoring Functions. The underlying idea of empirical scoring functions is that the binding free energy of a noncovalent receptor-ligand complex can be factorized into a sum of localized, chemically intuitive interactions. Such decompositions can be a useful tool to gain some insight into binding phenomena, even without analyzing 3D structures of receptor-ligand complexes. Andrews and colleagues derived average functional group contributions to the binding free energy by analyzing a set of 200 compounds for which the affinity to a receptor had been experimentally determined (266). Such average functional

group contributions can then be used to estimate the mean overall binding affinity of a compound independent of a particular binding site. This value can be compared to the experimental binding free energy: if the experimental affinity is similar to or even more favorable than the computed one, the ligand obviously shows a good fit with the receptor and its functional groups are supposedly all involved in interactions with the protein; on the other hand, if it is significantly less favorable, the compound apparently does not fully exploit its potential to form optimal interactions. Similarly, experimental binding affinities have been analyzed on a per-atom basis in quest of the maximal binding affinity of noncovalent ligands (288). It was concluded that in the strongest binding ligands each non-hydrogen atom on average contributes 6.3 kJ/mol to the binding energy.

The analysis of binding phenomena can be performed with much more detail if the 3D structures of receptor-ligand complexes are available. Based on the assumption of additivity, the binding affinity ΔG_{bind} can be estimated as a sum of interactions multiplied by weighting factors:

$$\Delta G_{\text{bind}} \approx \sum \Delta G_i f_i.$$

Here, each f_i corresponds to an interaction term that depends on structural features of the complex and each ΔG_i represents a weighting coefficient, which is determined on the basis of a training set of experimental affinities for crystallographically known protein–ligand complexes. Scoring schemes that use this concept are called empirical scoring functions. Several reviews summarize details of individual parameterizations (26, 44, 56, 289–292). The individual terms in empirical scoring functions are usually chosen such that they intuitively cover important contributions of the total binding free energy. Most empirical scoring functions are derived by evaluating the functions f_i on a set of protein–ligand complexes and fitting the coefficients ΔG_i to experimental binding affinities of these complexes by multiple linear regression or supervised learning. The relative weight of the individual contributions depends on the training set. Usually, between 50 and 100 complexes are used to derive the weighting factors. In a recent study it has been shown that many more than 100 complexes were necessary to achieve convergence (293). The reason for this finding is probably the fact that the publicly available protein–ligand complexes fall in a few rather strongly populated classes.

Empirical scoring functions usually contain individual terms for hydrogen bonds, ionic interactions, hydrophobic interactions, and binding entropy. Hydrogen bonds are often scored by simply counting the number of donor–acceptor pairs that fall into a given distance and angle range favorable for hydrogen bonding, weighted by penalty functions for deviations from ideal standard values (80, 294–296). The amount of error tolerance in these penalty functions is critical. If large deviations from the ideal are tolerated, the scoring function cannot discriminate sufficiently between different placements of a ligand, whereas too stringent tolerances artificially score similar complexes rather differently. Attempts have been described to reduce the strong distance dependency of such interactions by assigning soft modulating functions on an atom-pair basis (297). Other concepts try to avoid penalty functions and introduce distinct regression coefficients for strong, medium, and weak hydrogen bonds (293). The Agouron group has used a simple four-parameter potential that is a piecewise linear approximation of a potential neglecting angular terms ("PLP scoring function") (185). Most functions consider all types of hydrogen bonds equivalently. Some attempts have been made to distinguish between different donor–acceptor functional group pairs. Hydrogen bond scoring in GOLD (176, 177) is based on a list of hydrogen bond energies, derived from *ab initio* calculations, for any combination of 12 donor and 6 acceptor atom types. A similar differentiation of donor and acceptor groups is attempted in the program GRID (218) for the characterization of binding sites (219,220, 298). The consideration of such lookup tables in scoring functions might help to avoid false predictions originating from an oversimplification of some individual interactions.

Reducing the weight of hydrogen bonds localized at the solvent-exposed rim of a binding

site is a useful concept to avoid false positives in virtual screening. This is achieved by reducing charges of surface-exposed residues in cases where explicit electrostatic terms are used (274) or by multiplying the hydrogen bond contribution with a factor that depends on the accessibility of the involved protein counter group (299).

Ionic interactions are handled in a way similar to hydrogen bonds. Long-distance charge-charge interactions are usually neglected, and it is thus more appropriate to refer to salt bridges or charge-assisted hydrogen bonds. The scoring function by Boehm implemented in LUDI (294) assigns a stronger weight to salt bridges than to neutral hydrogen bonds. This differentiation generally proved successful in scoring series of thrombin inhibitors (295, 300). However, comparable to force field scoring, the danger exists that highly charged molecules receive overestimated scores. Experience with FlexX containing a variant of Boehm's scoring function has shown that more reliable predictions are obtained if charged and uncharged hydrogen bonds are handled equally in a virtual screening application. Similar experience has also been collected using the ChemScore function (80).

Hydrophobic interactions are usually calibrated to the size of the contact surface buried upon receptor-ligand complex formation. Often, a reasonable correlation between experimental binding energies can be achieved considering only a surface term [see, for example, (1, 301, 302) and the discussion in Section 2.1.]. Various approximations for such surface terms have been described, for example, as grid-based (294) or volume-based approaches (cf. the discussion in Ref. 115). Many functions are based on a distance-dependent summation over neighboring receptor-ligand atom pairs. Distance-dependent cutoffs have been introduced in various ways, either short (110) or longer to include atom pairs that are not involved in direct van der Waals contacts (80, 185). The weighting factor ΔG_i of the hydrophobic term depends strongly on the training set. Supposedly, this fact has been underestimated in the development of many empirical scoring functions (35) because in most training sets ligands composed of numerous donor and acceptor groups are overrepresented (many peptide and carbohydrate fragments).

In most empirical scoring functions, a hydrophobic character is attributed to several atom types, with equivalent weight for all hydrophobic contributions. In a more sophisticated approach, the propensity of particular atom types to be solvent-exposed or embedded in the interior of a protein can be assessed by so-called atomic solvation parameters. These have been derived, for example, from experimental octanol/water partition coefficients (303, 304) or from protein crystal structures (305, 306). Atomic solvation parameters are used in the VALIDATE scoring function (307) and have been tested in DOCK (308).

Entropy terms account for the restriction of conformational degrees of freedom of the ligand upon complex formation. A crude but useful estimate of this entropy contribution is the number of rotatable bonds of a ligand (294, 296). This measure has the advantage of being a function of the ligand only. More sophisticated estimates try to take into account the nature of the ligand portion on either side of a flexible bond, particularly with respect to the interactions formed with the receptor (80, 307). This concept is based on the assumption that purely hydrophobic contacts allow for more residual motion in the ligand fragments.

4.1.3 Knowledge-Based Methods. Empirical scoring functions regard only those interactions that are explicitly part of the model. Less frequent interactions are usually neglected, even though they *can* be strong and specific, for example, NH-π hydrogen bonds. To generate a comprehensive and consistent description of all these interactions in the framework of empirical scoring functions would be a difficult task. However, the exponentially growing body of structural data on receptor-ligand complexes *can* be exploited to discover favorable binding geometries. "Knowledge-based" scoring functions try to capture the knowledge about protein-ligand binding that is implicitly stored in the protein data bank by means of statistical analysis of structural data, without referring to often inconsistent experimentally determined binding affinities (309). They are based on the concept of the inverse formulation of the Boltzmann law,

$$E_{ij} = -kT \ln(p_{ijk}) + kT \ln(Z),$$

where the energy function E_{ij} is called a potential of mean force for a state defined by the variables i, j, and k; p_{ijk} is the corresponding probability density; and Z is the partition function. The second term of the sum is constant at constant temperature T and does not have to be regarded, given that $Z = 1$ can be selected by defining a suitable reference state, which leads to normalized probability densities ~~The inverse Boltzmann approach has been applied to assemble potentials from databases of protein structures to score protein models in the context of protein structure prediction (310). To establish a function to score protein–ligand complexes, the variables i, j, and k are assigned to address protein and ligand atom types, and their interatomic distances. The occurrence frequency of individual contacts is a measure of their energetic contribution to binding. If a specific contact occurs more frequently than expected by random or seen in an average distribution, it is assumed to be favorable. On the other hand, if it occurs less frequently, repulsive or unfavorable interaction between two atom types is anticipated. The frequencies are thus converted into sets of atom-pair potentials ready for further evaluation.

First applications in drug research (134, 311,312) were restricted to small data sets of HIV protease-inhibitor complexes and did not result in generally applicable scoring functions. Recent publications (226, 313–318), however, have shown the usefulness of these approaches. The first general-purpose function using such potentials was implemented in the de *novo* design program SMoG (313,314).

The PMF function by Muegge and Martin (317), consists of a set of distance-dependent atom-pair potentials $E_{ij}(r)$ that are expressed as

$$E_{ij}(r) = -kT \ln[f_j(r)\rho^{ij}(r)/\rho^{ij}].$$

Here, r is the atom pair distance, and $\rho^{ij}(r)$ is the number density of pairs ij in a certain radius interval about r. This density is calculated by the following procedure. First, a maximum search radius is defined. This radius describes a reference sphere around each ligand atom j. Receptor atoms of type i are searched within this sphere. Subsequently, the sphere is subdivided into shells of a predefined thickness. The number of receptor atoms i matching each spherical shell is divided by the volume of this shell and averaged over all occurrences of ligand atoms j in the evaluated data set of protein–ligand complexes. The term ρ^{ij} in the denominator is the average density of receptor atoms i falling into the whole reference volume. It is argued that the spherical reference volume around each ligand atom needs to be corrected by eliminating the occupied volume of the ligand itself, given that ligand-ligand interactions are not regarded in this area. This is done by a volume correction factor $f_j(r)$ that is a function of the ligand atom only and gives a rough estimate of the preference of an atom of type j to be exposed rather than buried in the ligand. Muegge could show that the volume-correction factor contributes significantly to the predictive power of the PMF function (319). Also, reference radii between 7 and 12 Å are applied to implicitly include solvation effects, especially the propensity of individual atom types to be located inside a protein cavity or in contact with solvent (320). To rank docking solutions, the PMF function is evaluated in a grid-based fashion and combined with a repulsive van der Waals potential at short distances.

The DrugScore function by Gohlke et al. (226) is based on roughly the same formalism, albeit with several differences in the derivation that lead to different potential forms. Most notably, the statistical distance distributions $\rho^{ij}(r)/\rho^{ij}$ for the individual atom pairs ij are divided by a common reference state that is taken as the average over the distance distributions of all atom pairs $\rho(r) = \Sigma\Sigma\ \rho^{ij}(r)/ij$. To consider only direct ligand-protein contacts, the upper sample radius has been set to 6 Å. At this distance, no further atoms (e.g., of a water molecule) can mediate a protein–ligand interaction. The individual potentials have the form

$$E_{ij}(r) = -kT(\ln[\rho^{ij}(r)/\rho^{ij}] - \ln[\rho(r)]).$$

These pair potentials are used in combination with potentials depending on single (protein or ligand) atom types that express the propensity of an atom type to be buried within a particular

protein environment on complex formation. Contributions of these surface potentials and the pair potentials are weighted equally in the final scoring function. This scoring function has initially been developed with the primary goal to differentiate between correctly docked (near native) ligand poses versus decoy binding modes for the same protein-ligand pair. However, through appropriate scaling also quantitative estimates across different protein–ligand complexes are possible (318).

Mitchell and coworkers choose a different type of reference state for their BLEEP potential (315). The pair interaction energy is written as

$$E_{ij}(r) = kT \ln[1 + m^{ij}\sigma] - kT \ln[1 + m^{ij}\sigma\rho^{ij}(r)/\rho(r)].$$

Here, the number density $\rho^{ij}(r)$ is defined as above, but it is normalized by the occurrence frequency of all atom pairs at this same distance instead of by the number of pairs ij in the whole reference volume. The variable m^{ij} is the number of pairs ij found in the evaluated data set, and σ is an empirical factor that defines the weight of each observation. This potential is combined with a van der Waals potential as a reference state to compensate for the lack of sampling at short distances and for certain underrepresented atom pairs.

Besides differences in the functional form and reference state, from a more practical point of view, the knowledge-based potentials differ also with respect to scope of atom type definitions and the amount of structural data used for their derivation. The number of different atom types ranges from 17 in DrugScore to 40 nonmetal atom types in BLEEP. In all cases, the Protein Data Bank (321) was the source of the solved crystal structures. For BLEEP 351 selected complexes were used, whereas the PMF function was extracted from 697 complexes, and DrugScore was derived using 1376 complexes. In the latter case, the data have been extracted from Relibase (322,323).

4.2 Critical Assessment of Current Scoring Functions

4.2.1 Influence of the Training Data. All fast scoring functions share a number of deficiencies that one should be aware of in any application. First, most scoring functions are in some way fitted to or derived from experimental data. The functions necessarily reflect the accuracy of the data that were used for their derivation. For instance, a general problem with empirical scoring functions is the fact that the experimental binding energies usually originate from many different sources and therefore consist of a rather heterogeneous data set affected by all kinds of experimental errors. Furthermore, scoring functions mirror not only the quality but also the scope of experimental data used for their development. Virtually all scoring functions are still derived from data mostly based on high-affinity receptor-ligand complexes. Many of these are still of peptidic nature, whereas interesting leads in pharmaceutical research are non-peptidic. This is reflected in the relatively high contributions of hydrogen bonds in the total score. The balance between hydrogen bonding and hydrophobic interactions is a very critical issue in scoring, and its consequences are especially obvious in virtual screening applications, as illustrated in Section 5.3.

4.2.2 Molecular Size. The simple additive nature of most fast scoring functions often leads to gradually increasing scores for molecules of larger size. Although it is true that small molecules with a molecular weight below 200–250 rarely show very high affinity, there is no physical reason why larger compounds should automatically possess higher activity. Comparing the scores of two compounds of significant size difference therefore calls for a term that compensates the size dependency. In some applications, a constant "penalty" term has been added to the score for each heavy atom (324) or a term proportional to the molecular weight has been considered (325). The empirical scoring function implemented in the docking program FLOG has been normalized to remove the linear dependency of the crude score on the number of ligand atoms (121). Originally introduced to improve the correlation between experimental and calculated affinities, entropy terms reflecting the change in conformational mobility upon ligand binding also help to reduce an excessive score for overly large and flexible mol-

ecules (80,294). The size of the solvent-accessible surface of the ligand in its bound state can also be used as penalty term to discard large ligands not fully buried in the binding site. It should be noted, however, that all these approaches are very pragmatic in nature and do not solve the problem of size dependency, which is closely related to a proper understanding of cooperativity effects (265).

4.2.3 Penalty Terms. In general, scoring functions reward favorable interactions such as hydrogen bonds, but rarely penalize unfavorable ones. They are derived from experimentally determined crystal structures, and thus nonnative and energetically unfavorable orientations of a ligand within the binding site are not observed and can hardly be accounted for in a regression-based scoring function. Knowledge-based scoring functions try to capture such effects by referring to a reference state that corresponds to a mean situation. At first glance, the neglect of angular terms in the compilation of knowledge-based scoring functions results in averaged pair potentials that may not discriminate sufficiently between different binding geometries. However, some degree of angular dependency is considered, given that pair potentials for different atom types are always evaluated in combination with each other (226). Obvious deficiencies in regression-based scoring functions, such as electrostatic repulsions and steric clashes, can be avoided by defining reasonable penalty terms or by importing them from molecular mechanics force fields. This has been realized in the "chemical scoring" function implemented in the docking program DOCK (106, 126, 137, 273, 326, 327), which is a modified van der Waals potential being attractive or repulsive between individual groups of donor, acceptor, and lipophilic receptor and ligand atoms (224, 328). Other situations cannot be avoided by simple "clash" terms, but require a more sophisticated analysis of binding geometry. Among these are incomplete steric filling of the binding cavity by a ligand within the cavity, an unreasonably large ligand surface area remaining solvent exposed in the complex or the formation of voids at the receptor-ligand interface. Possible approaches to resolve these shortcomings are empirical filters that detect such unsatisfactory solutions and remove them according to user-specified thresholds (329). A promising approach to properly reflect such cases is the inclusion of artificially generated, erroneous, decoy solutions in the optimization of scoring functions as reported for the scoring function of a flexible ligand superposition algorithm (330,331).

4.2.4 Specific Attractive Interactions. Another general deficiency of scoring functions is the simplified description of attractive interactions. Molecular recognition is not entirely based on hydrogen bonding and hydrophobic contacts. Especially in host-guest chemistry, other specific types of interactions are frequently used to characterize the observed phenomena. For example, hydrogen bonds are formed between acidic protons and π-systems (332). These bonds can substitute for conventional hydrogen bonds in strength and specificity, as has been noted in protein-DNA recognition (333). Another type of less frequently observed interactions is the cation-T interaction, which is especially important at the surface of proteins (39, 334). Current empirical scoring functions usually neglect these interactions. Similarly, the directionality of interactions between aromatic rings is hardly considered (335, 336). Because of the regression-type adjustment, some energy contributions originating from these interactions are already implicitly incorporated into the conventional interaction terms. This might be one explanation why hydrogen bond contributions are traditionally overestimated in regression-based scoring functions. Knowledge-based approaches automatically incorporate these interactions in a scoring function, provided they occur with reasonable frequency in the data set used to develop the potentials.

4.2.5 Water Structure and Protonation State. Uncertainties about protonation states and the involvement of water in ligand binding further complicate scoring. These considerations are relevant for the development as well as the application of scoring functions. As mentioned above, the entropic and enthalpic contributions involving the reorganization of water molecules upon ligand binding are very difficult to predict (see, for example, Ref. 337). Currently, the most pragmatic approach to handle the water problem is the elucidation of

"conserved water molecules" and to consider them as part of the receptor. A knowledge-based tool to estimate the "conservation" of water molecules upon ligand binding has been developed (217) and incorporated into a docking procedure (111) (cf. Section 3.2.2). It is based on crystallographic information and tries to extract rules about water sites by analyzing whether they are recurrently occupied by water molecules in series of related protein–ligand complexes.

Scoring functions require predefined atom types for each protein and ligand atom. This also implies the fixed assignment of a protonation state to each acidic and basic group. Knowledge-based functions, which do not consider hydrogen atoms, are equally affected by the problem because the atom type definitions normally imply a certain protonation state. Presently, such estimates might be reliable enough for the situation in aqueous solution; however, significant pK_a shifts are possible upon ligand binding (338) as a result of strong changes of the local dielectric conditions. They give rise to protonation reactions in parallel to the binding process. With respect to scoring, switching from a donor to an acceptor functionality because of altered protonation states has important consequences (279). Accordingly, improved docking and scoring algorithms must incorporate a more detailed and flexible description of protonation states.

4.2.6 Performance in Structure Prediction and Rank Ordering of Related Ligands. Similar to the broad range of available docking tools (cf. Section 3.2.3), the multitude of different scoring schemes calls for an objective assessment to evaluate their scope and limitations. This depends in part on the anticipated application; that is, whether protein–ligand complexes should be predicted (using the scoring scheme as objective function in docking), whether a set of ligands should be ranked with respect to one target protein (K_i prediction), or whether the scoring function is used to select possible hits out of a large database of candidate molecules (virtual screening).

An objective assessment of the available scoring functions is difficult because only very few functions have been tested on the same data sets or with the same docking tool. For structure prediction, several studies have shown that knowledge-based scoring functions are at least equivalent to regression-based functions. The PMF function has been successfully applied to structure prediction of inhibitors of neuraminidase (339) and MMP3 (229) in combination with the program DOCK, yielding superior results to the DOCK force field and chemical scoring. The DrugScore function was tested on a large set of PDB complexes and gave significantly better results than those of the original FlexX scoring function using solutions generated by FlexX as the docking engine. DrugScore performed similarly to the force field score in DOCK, but outperformed the chemical scoring (226). Moreover, with respect to the correlation between experimental and calculated binding energies, very promising results have been obtained with DrugScore (318) and PMF (229, 317, 319, 339). BLEEP has recently been tested for scoring docked protein–ligand complexes (340). It was found to be slightly better than the DOCK energy function in discriminating decoy situations from near-native binding modes.

Although in many docking programs the same function is applied as an objective function for structure generation and for energy evaluation, better results can sometimes be obtained if different functions are applied. In particular, the docking objective function can be adapted to the docking algorithm used. In a parameter study, Vieth et al. found that using a soft-core van der Waals potential made their MD-based docking algorithm more efficient (274). Using FlexX as the docking engine, we observed that the original FlexX scoring function emphasizes directional interactions (mostly hydrogen bonds) in the docking phase. Subsequently, the ranking of individual library entries can be done successfully with a simple PLP potential that lacks directional terms, but considers general steric fit of receptor and ligand. Results are significantly worse if PLP is used already in the incremental built-up procedure of the docked ligand.

It is even more difficult to draw valid conclusions about the relative performance of scoring functions to rank sets of inhibitors with respect to their binding affinity for the same target. First, there is hardly any pub-

lished study in which different functions have been applied to the same data sets. Second, experimental data are often not measured under the same conditions but collected from various literature references. This retrieval from various sources usually implies larger uncertainties within the experimental data

The task of ranking sets of 10–100 related ligands with respect to one target can also be handled by computationally more demanding methods. The most general approaches are probably force field scores complemented by electrostatic desolvation and surface area terms. An example is the MM-PBSA method that combines Poisson-Boltzmann electrostatics with AMBER molecular mechanics calculations and MD simulations (341,342). This method has recently been applied to an increasing number of examples, showing quite promising results (343–346). Poisson-Boltzmann calculations have been performed on a variety of targets with many related computational protocols (280–282, 347–350). Alternatively, extended linear response protocols (263) can be used. The OWFEG grid method by Pearlman has also shown very promising

VIRTUAL SCREENING

As outlined in section, virtual screening is a multistep process. Although, in principle, the whole process can be fully automated, it is highly advisable to allow for manual interventions, in that visual inspection and selection still play a major role.

The process usually starts with a detailed analysis of the available 3D protein structures. If possible, highly homologous structureswill also be analyzed, either to generate additional ideas about possible ligand structural motifs or to gain some insight on how to achieve selectivity against other proteins of the same class. A superposition of different protein–ligand complexes provides some ideas about key interactions repeatedly found in tight-binding protein–ligand complexes. Such an overlay will also highlight flexible parts of the protein or recurring water molecules in the binding site that could be included in the docking process. Tools such as Relibase (322, 323) may be used to perform these comparative analyses of protein–ligand complexes in an efficient way. Subsequently, programs like GRID (218), LUDI (108, 109), Superstar (351, 352), or DrugScore (318) are used to visualize potential binding sites ("hot spots") in the active site; in principle, any scoring function could be used for this purpose.

An important result of the 3D structure analysis is usually the identification of one or more key interactions that all ligands should satisfy. In aspartic proteases, for example, inhibitors should form at least one hydrogen bond to the catalytic Asp side-chains, whereas in metalloproteinases a coordination to the metal seems mandatory. Sometimes, a known ligand portion is used as initial scaffold based on which virtual screening techniques search for optimal side-chains. In principle, this step is not required, and instead one could fully rely on the docking and scoring step. However, following a pragmatic approach, it is important to use any well-founded information that is available about the system under consideration because more valuable results can usually be expected this way.

Once a reasonable hypothesis about the binding-site requirements has been generated, the next level of virtual screening is approached. Whether databases of commercially available compounds or "virtual" libraries of designed compounds are screened, it is advisable not to dock every possible compound, but only those that pass a series of hierarchical filters (cf. also Fig. 7.3). Simple preliminary filters remove

- compounds with reactive groups such as $-SO_2Cl$ or $-CHO$ because they are expected to cause problems in some biological assays as a result of unspecific covalent binding to the protein.
- compounds with molecular weights below 150 or above 500. Small molecules such as benzene are known to bind to proteins rather unspecifically at several sites. Large molecules such as polypeptides are difficult to optimize subsequently, given that good

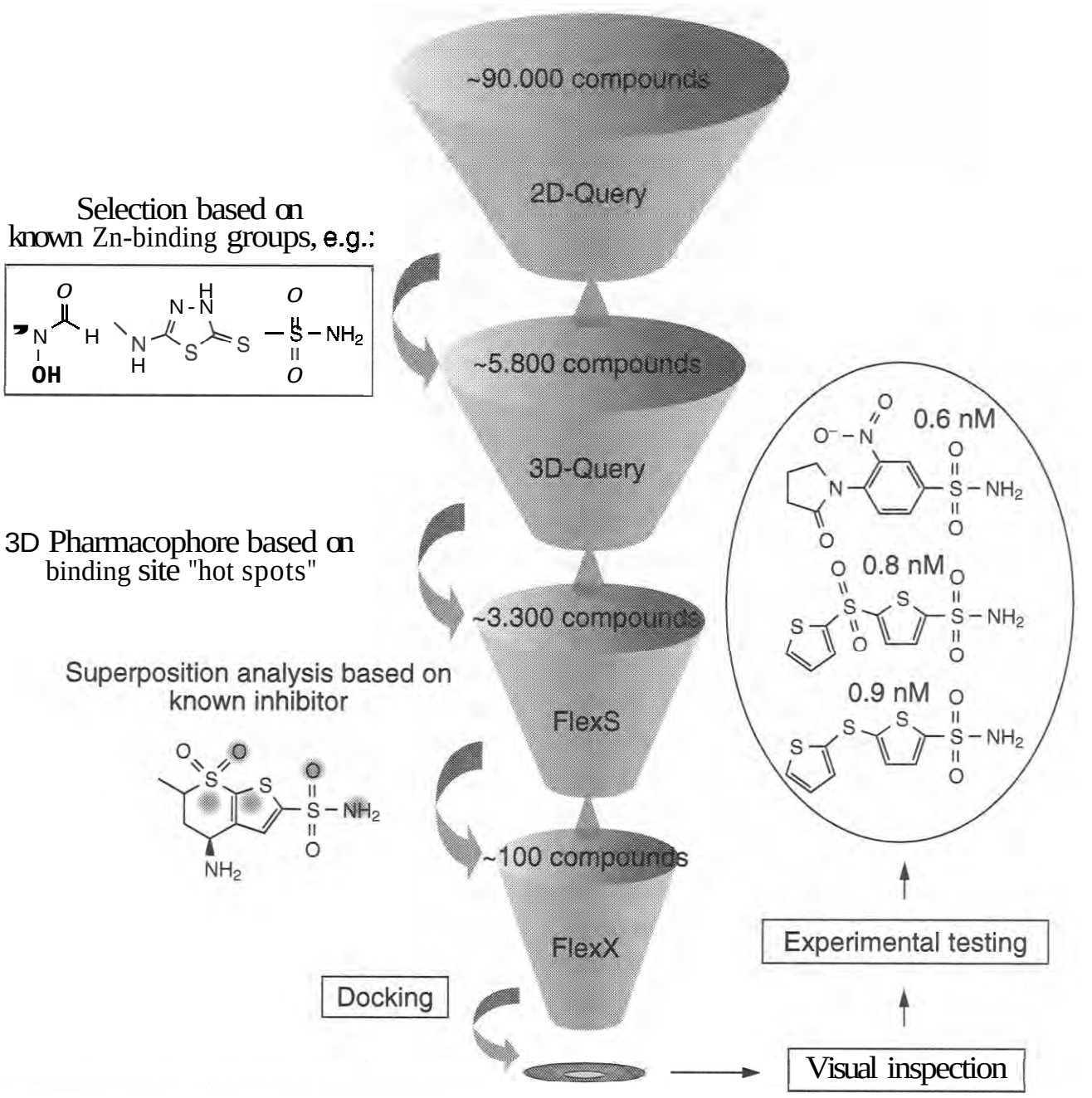


Figure 7.3. Hierarchical filtering process in virtual screening for carbonic anhydrase inhibitors.

bioavailability is usually limited to compounds with molecular weights below 500. compounds termed as "non-drug-like" according to criteria extracted from known drugs (353,354).

After this general preselection, it can be advantageous to apply further steps of hierarchical filtering. As mentioned above, this could involve the selection of functional groups inevitably required to anchor a ligand to the most prominent interaction sites. Subsequently, the information of the "hot spot" analysis—translated into a pharmacophore hypothesis—can be used as matching criterion for a fast database screen. Such tools either involve fast tweak searching (355)or scan over precalculated conformers of the candidate molecules (356). The list of prospective hits can then be submitted to a similarity search using information about already known active compounds, which could either be ligands already structurally characterized by crystallography or hits from a complementary HTS study. This optional step of the analysis tries to incorporate all available information about known hits and produces a reranking of the candidate molecules to be submitted to docking. As tools for the spatial similarity analysis, Feature Trees (357, 358), FlexS (330), and SEAL (254–256) have all been successfully applied.

All remaining ligands are docked into the target protein and a list of some hundred to several thousand small molecules, each with a computed score, is produced. These have to be further analyzed to discard undesirable structures. Selection criteria could be any of the following:

- Lipophilicity (if not addressed before). Highly lipophilic molecules are difficult to test because of their low solubility.
- Structural class. If 50% of the docked structures belong to one single chemical class, it is probably not necessary to test all of them (359).
- Unreasonable docked binding mode. Fast docking tools cannot produce reliable solutions for all compounds; often, even some well-scoring compounds are simply docked to the outer surface of the protein or adopt rather strained conformations to achieve good surface complementarity within the binding pocket. Computational filters help to detect such situations (329).

Finally, the selected compounds are ordered or synthesized and then tested. If the goal is to identify even weakly binding ligands as first leads, sufficient sensitivity of the biological assay has to be ensured [cf., for example, Ref. 161. In this context it has also to be considered that limited solubility of the hits in water or water/DMSO mixtures often hampers affinity determinations at high concentrations.

Successful virtual screening has to produce a set of compounds significantly enriched with active compounds compared to random selection. A key parameter to assess the performance of docking and scoring in virtual screening is therefore, at least in theoretical case studies, the so-called enrichment factor. It is simply the ratio of active compounds in the subset selected by docking divided by the number of active compounds found in a randomly selected subset of equal size. To record such enrichment factors also for controlling performance at the various filter steps, a set of known active compounds is mixed with the set of candidate molecules. This strategy, however, requires a set of reasonable size (e.g., 30–50 ligands), which is not always given in a real-life virtual screening study. Furthermore, enrichment factors are far from being ideal indicators, particularly at later filter steps where a (hopefully) increasing amount of active compounds detected among the entries of the database competes with the set of known active ligands and artificially lowers the enrichment factor.

5.1 Combinatorial Docking

Docking of large compound collections requires fast algorithms. If the collection is an unstructured library of more or less unrelated compounds, each individual molecule must be docked independently ("sequential docking"), and only the fastest docking methods are applicable in this context, unless massive computer resources are used, as in the DockCrunch project based on the PRO–LEADS program (360). Examples for such fast docking tools are SLIDE (111) or the docking method by Diller and Merz (112). Both have been developed for database screening and library prioritization. Before docking, it is generally advisable to eliminate compounds that would provide only redundant information (similarity filters) or are very unlikely to yield high scores. Clearly, the filter routines need to be faster than the docking and scoring procedure, but this is normally the case.

Complementary to initial filtering, a preorganization of compounds into families exhibiting some kind of similarity has been demonstrated to improve the results of database screening. In the strategy shown by Su et al. (359), all molecules of any family are docked and scored, but only the best-scoring member of a high-ranking family is allowed to remain in the final hit list, whereas the scores of related molecules are recorded as annotations to this representative family member. This increases the diversity of the hit list and helps to identify a higher number of different classes of potential ligands.

An alternative to sequential docking can be followed if combinatorial libraries are evaluated. Quite a few programs have been specifically designed for speed-up by so-called combinatorial docking. They profit from the structured, incremental nature of combinatorial libraries and the fact that molecules of a combinatorial library consist of a common core. This core is assumed to form common specific interactions with the receptor (possibly supported by experimental evidence) and can thus be prepositioned in the binding pocket in one or a few similar orientations. It then serves as skeleton for the addition of substituents. Obviously, this step is ideally suited for incremental construction algorithms (361)

and significantly reduces the complexity of the docking problem, limiting the required computation time per ligand. Earlier examples of this combinatorial docking approach are PRO_SELECT (362) and CombiDOCK (324). The latter is based on the DOCK program and has recently been enhanced by a vector-based orientation filter, to ensure productive scaffold poses, and by a free-energy-based scoring procedure (279). Another recent combinatorial docking procedure has been implemented as FlexX^c extension in FlexX (363). It follows a recursive scheme to traverse the combinatorial library space efficiently. The algorithm is based on a tree data structure that allows the efficient reuse of previously calculated docking results. FlexX^c follows the library search tree in a depth-first manner, whereas CombiDOCK uses a breadth-first approach to evaluate fragments attached to a scaffold. A general advantage of breadth-first searches is that they allow for an efficient pruning of the search tree based on the scoring values.

De novo design tools have also been adapted to the problem of combinatorial docking and combinatorial library design. The program LUDI, for example, has been enhanced by the ability to connect building blocks in a chemically and structurally adequate manner; it can thus be used for combinatorial docking by fitting building blocks onto the interaction sites and simultaneous linking to previously docked core fragments (300). It has been successfully applied in the design of new thrombin inhibitors accessible through a single reaction. Another example is a variant of the Builder program (364) that was used to select substituents for a library of cathepsin D inhibitors (12). Yet another approach is DREAM++, a suite of programs for the design of virtual combinatorial libraries (365). Here, the DOCK algorithm is used for the molecular placement. Variable fragments are joined consecutively in compliance with predefined types of well-characterized organic reactions. Speed-up is achieved by preserving ("inheriting") information about common partial structures across different reactions, such that only the conformations of newly added fragments are searched.

Generally speaking, combinatorial docking approaches work best in cases where a core fragment plays a dominant role in the binding process and can be placed with high confidence in a well-characterized specificity pocket, such as the S1 pocket in thrombin. A further issue to consider is mutual fragment dependencies, that is, when multiple fragments are hooked up to a scaffold in a sequential manner; the results can depend on the sequence by which they are added (see, for example, Ref. 363). Thus, in unfavorable cases, different orders of attachment have to be followed to circumvent this possible limitation.

5.2 Seeding Experiments to Assess Docking and Scoring in Virtual Screening

True enrichment factors can be calculated only if experimental data are available for the full library, although such situations are unusual. Accordingly, studies using enrichment factors as a figure-of-merit to assess the performance of a virtual screening can serve for theoretical validation purposes only. Several authors have tested the predictive ability of docking and scoring tools by compiling an arbitrary set of diverse, drug-like compounds complemented by a number of known active compounds. This "seeded" library is then subjected to the virtual screening, and for the purpose of assessment it is assumed that the added active compounds are the only true actives in the library. Clearly, this is a rather questionable assumption.

Several seeding experiments have been published. An example has been performed at Merck using FLOG (121). A library consisting of 10,000 compounds including inhibitors of various types of proteases and HIV protease was docked into the active site of HIV protease. This resulted in excellent enrichment of the HIV protease inhibitors: all inhibitors but one were among the top 500 library members. However, inhibitors of other proteases were also considerably enriched (366).

Seeding experiments also allow for comparisons of different docking and scoring procedures, as shown, for example, by Charifson et al. (86), Bissantz et al. (230), and Stahl and Rarey (287). Charifson et al. compiled sets of several hundred active molecules for three different targets, p38 MAP kinase, inosine monophosphate dehydrogenase, and HIV protease. These were docked into the corresponding active sites together with 10,000 randomly se-

lected, but drug-like, commercially available compounds using DOCK (327) and the Vertex in-house tool Gambler. ChemScore (80, 188), the DOCK AMBER force field score, and PLP (185) performed consistently well in enriching active compounds. This result was partially attributed to the fact that a rigid-body optimization could be carried out with these functions because they include repulsive terms in contrast to many other tested functions. Stahl and Rarey compared DrugScore (226), PMF (317), PLP (185), and the original FlexX score using FlexX for docking (110,130,138). Interestingly, the two knowledge-based scoring functions performed differently. DrugScore achieved better ranking for the tight-binding ligands in narrow lipophilic cavities of COX-2 and the thrombin S1 pocket. In contrast, PMF obtained better enrichment for the case of the very polar binding site of neuraminidase. Obviously, a general strength of PMF is the description of complexes showing multiple hydrogen bonds. This has also been noted in the study by Bissantz et al., in which PMF was found to perform well for the polar target thymidine kinase and less well for the estrogen receptor (230).

5.3 Hydrogen Bonding versus Hydrophobic Interactions

A balanced description of the contribution of hydrogen bonding and hydrophobic interactions to the total score is of general importance, to avoid a bias toward either highly polar or completely hydrophobic molecules. The actual parameterization of a scoring function depends on the compilation of the data set used to develop the function. Empirical scoring functions are more likely affected by the data set composition used for parameterization, but can be quickly reparameterized. In the case of knowledge-based functions such a readjustment is more difficult to perform; however, because of the much larger databases used for their development, they are supposed to be less dependent on special data set compilations.

The PLP function, for example, addresses general steric complementarity and hydrophobic interactions based on rather long-range pair potentials, whereas the FlexX score concentrates on hydrogen-bond complementarity. This is clearly reflected in results of database-ranking experiments. To combine the virtues of both scoring functions and to construct a more robust general function, a combination of PLP and FlexX called ScreenScore has recently been published (287). It was derived by a systematic optimization of library ranking results over seven targets and covers a wide range of active sites with respect to form, size, and polarity. ScreenScore obtains good enrichments for COX-2 (highly lipophilic binding site) and neurarninidase (highly polar site), whereas the individual functions fail in one of the two cases. The authors of PLP have recently enhanced their scoring function by including directed hydrogen bonding terms (367). Similar to ScreenScore, this could also lead to a more robust scoring function.

5.4 Finding Weak Inhibitors

Seeding experiments are often carried out with a small number of active compounds that are already optimized for binding to the studied target. Enrichment factors based on the retrieval of these compounds are not very conclusive because the recovery of potent inhibitors from a large set of candidate molecules is significantly easier than the discovery of new, but usually rather weak inhibitors from a large majority of nonbinders. In general, as in HTS, one can only expect hits from virtual screening that bind in the low micromolar range.

Nevertheless, a recent study showed that library screening can also successfully detect very weak ligands. Approximately 4000 commercially available compounds had been screened for FKBP-binding by means of the SAR-by-NMR technique (368) and 31 compounds with activity in the low millimolar range were detected. This set of compounds was flexibly docked into the FKBP binding site using DOCK 4.0 with the PMF scoring function (369). Interestingly, significant enrichment factors of 2 to 3 were achieved, whereas scoring with the standard AMBER score of DOCK did not really provide an enrichment.

5.5 Consensus Scoring

Different scoring schemes focus on different aspects as most important contributions to

binding. However, these differences do not necessarily become obvious when calculating binding affinities of known active compounds. In contrast, the scoring of non-active compounds could unravel such differences. Vertex has reported good experience with so-called consensus scoring. Here, docking results are scored by several distinct functions and only those hits are considered that are rendered prominent by several of the functions. A significant decrease in false positives has been described (86), but inevitably a number of true positives is lost (see, for example, Ref. 230).

When consensus scoring is applied, one should thus keep in mind that, although the number of false positives can be reduced, the danger exists to discard some active compounds highlighted by only one of the scoring functions. This would, for example, apply to the above-mentioned PLP and FlexX scoring functions, which emphasize different aspects of ligand binding. Here, consensus scoring could be counterproductive. Therefore, along with consensus scoring, the individual scoring results should be consulted. Generally, however, it appears that one can expect more robust results from consensus scoring.

5.6 Successful Identification of Novel Leads through Virtual Screening

A considerable number of publications have proved that virtual screening can be efficiently used to discover novel leads (11, 13, 142, 370–375). Some of the most recent examples are briefly presented in the following.

The program ICM has been used to identify novel antagonists for a nuclear hormone receptor (201) and, together with DOCK, to find inhibitors for the RNA transactivation response element (TAR) of HIV-1 (25). The virtual screening protocol started with 153,000 compounds from the Available Chemicals Directory (ACD) (376) and involved increasingly elaborate docking and scoring schemes as the screening proceeded toward smaller selections of compounds. In the HIV-1 TAR study, the ACD library was first rigidly docked into the binding site using the DOCK program along with a simple contact scoring scheme. Then, 20% of the best-scoring compounds were subjected to flexible docking with ICM and an empirical scoring function specifically tailored to RNA targets, providing a selection of approximately 5000 compounds. This was followed by two additional steps involving longer sampling of conformational space to retrieve 350 most promising candidates. Of these, a very small fraction was tested experimentally and two compounds were found to significantly reduce the binding of the Tat protein to HIV-1 TAR ($CD_{50} \sim 1\ \mu M$).

Recently, Grueneberg et al. discovered subnanomolar inhibitors of carbonic anhydrase II by virtual screening (15). The study was performed following a protocol of several consecutive steps of hierarchical filtering (Fig. 7.3). Carbonic anhydrase II is a metalloenzyme used as prominent target for the treatment of glaucoma. Its binding site is a rather rigid, funnel-shaped pocket. Known inhibitors such as dorzolamide bind to the catalytic zinc ion by a sulfonamide group. In a recent crystallographic study it could be demonstrated that only the sulfonamide group represents an ideal anchor for zinc coordination (377). An initial data set of 90,000 entries from the Maybridge (378) and LeadQuest (379) libraries was converted to 3D structures with Corina (380). In a first filtering step, compounds were requested to possess a known zinc-binding group. These compounds were then processed through UNITY (355) using a protein-derived pharmacophore query. The pharmacophore hypothesis had been constructed from a "hot spot" analysis of the available X-ray structures of the enzyme. This yielded a set of 3314 compounds. In a subsequent filtering step, the known inhibitor dorzolamide was used as a template onto which all potential candidates were flexibly superimposed by means of the program FlexS (330). The top-ranking compounds from this step were docked into the binding site with FlexX (110, 130, 138), taking into account four conserved water molecules in the active site. After visual inspection, 13 top-ranking hits were selected for experimental testing. Nine of these compounds showed activities below $1\ \mu M$, and three had K_i values below $1\ nM$. Two of the hits were also examined crystallographically. The docking solution predicted as best by DrugScore was found to be closer to the experimental structure than the one predicted by the FlexX score.

This strategy of hierarchical filtering starting with a mapping of candidate molecules onto a binding site-derived pharmacophore, followed by a similarity analysis with known ligands using either FlexS (3301, SEAL (254–256), or FeatureTrees (357, 358); and concluded by flexible docking with FlexX, which meanwhile was applied to three other proteins in the same laboratory. For t-RNA guanine transglycosylase, thermolysin, and aldose reductase, novel micromolar to submicromolar lead structures could be discovered. Most challenging in this context is aldose reductase because it performs pronounced induced fit changes upon ligand binding. Crystal structure analysis of a micromolar hit retrieved by virtual screening clearly revealed known and new areas of induced fit adaptation. The crystal structure obtained with this hit provides a good starting point for further lead optimization.

The de *novo* design of inhibitors of the bacterial enzyme DNA gyrase, a well-established antibacterial target (381), is another example for successful structure-based virtual screening, reported by Roche (16). HTS performed on the proprietary compound library provided no suitable lead structures. Therefore, a new rational approach was developed to discover potential lead structures using structural information of the ATP binding site in subunit B of the enzyme. At the onset of the project, the crystal structures of DNA gyrase subunit B complexed with a substrate analog and two inhibitors were available. In the buried part of the pocket they all donate a hydrogen bond to an aspartic acid side-chain and accept one from a conserved water molecule. As a design concept, the formation of these two key hydrogen bonds has been defined as mandatory. As an additional requirement, a lipophilic portion forming hydrophobic interactions with the enzyme was demanded. A new assay was established to allow for the detection of weakly binding inhibitors. A computational search of the ACD (376) and the Roche Compound Inventory identified hits having low molecular weights and matching the above-mentioned criteria. Relying on the results of the in *silico* screening Based on docking with LUDI and a pharmacophore search with CATALYST (356)] 600 compounds were tested initially. Then, close analogs of the first series of hits were assayed, resulting in a total screen of 3000 compounds. This provided 150 hits, clustered into 14 chemical classes. Seven of these classes could be demonstrated as novel DNA gyrase inhibitors competing for the ATP binding site. Subsequent structure-based optimization resulted in inhibitors with potencies equal to or up to 10 times better than those of known antibiotics.

6 OUTLOOK

The first docking programs were introduced about 20 years ago, and the publication of the first generally applicable scoring functions dates back about 10 years. Since then, much experience has been gained in developing and applying docking algorithms, using scoring functions, and assessing their accuracy. Significant progress has been made over the last few years and it appears as if there are now docking tools available to address a variety of goals with considerable accuracy, from the precise and detailed analysis of binding interactions for a small set of ligands up to a fast screening of large compound collections. Similarly, scoring functions are currently available that can be applied to a wide range of different proteins and consistently yield a considerable retrieval of active compounds. As a consequence, the pharmaceutical industry increasingly uses virtual screening to identify possible leads.

In fact, structure-based design is now established as an important approach to drug discovery complementing HTS (3821, although HTS has a number of serious disadvantages. It is expensive (383) and it leads to many false positives and a disappointingly small number of real leads (384, 3851, particularly if screening is performed on a member of a new protein class. Also, not all assays are easily amenable to HTS requirements. Finally, despite the library sizes of several million entries available to the pharmaceutical industry, these compound collections do not approach the size and diversity needed to even approximately cover the chemical space of drug-like organic molecules. Accordingly, focused design of novel compounds and com-

pound libraries should only gain importance. In light of current trends in structural genomics and patenting strategies, one may speculate that structure-based de *novo* design will become much more important in the near future.

To meet the increasing demands being placed on virtual screening, the development of more reliable scoring functions is certainly vital for success. In addition, novel or improved docking algorithms are required. We conclude by summarizing our perspective on major challenges in the further development of docking procedures and scoring functions:

1. The fact that protein–ligand interactions occur in aqueous solution is generally appreciated, but not yet adequately accounted for in molecular docking procedures. In particular, the simultaneous placement of explicit water molecules upon docking, accurate estimates of the water versus ligand interaction-energy balance, and the fast prediction of protonation states in binding pockets await a more satisfactory solution.
2. The consideration of a sufficient degree of protein flexibility needs to become part of standard docking approaches. This will require faster algorithms. In addition, with respect to scoring, an often overlooked aspect of this problem is that as soon as receptor flexibility is allowed, protein conformational energy changes need to be accounted for appropriately.
3. Although flexible-ligand docking has already become standard practice, the error rate in predictions of interaction geometries is still significant for more flexible ligands. Again, more efficient algorithms will be required to sample the conformation space more thoroughly.
4. Polar interactions are still not treated adequately. It is striking that, even though the role of hydrogen bonds in biology has been appreciated for a long time and the nature of hydrogen bonds is qualitatively well understood, their quantitative energetic description in protein–ligand interactions is still unsatisfactory (65).
5. All scoring functions are essentially expressed as simple analytical functions fitted to experimental binding data. The presently available crystal data on complex structures are strongly biased toward peptidic ligands. Because these data are used for the development of scoring functions, many overestimate the role of polar interactions. The development of improved scoring functions clearly requires access to better data, especially for nonpeptidic, low molecular weight, drug-like ligands, including weakly binding compounds.
6. Unfavorable interactions and unlikely docking modes are not penalized strongly enough. Methods for taking such undesired features into account are still lacking in presently available scoring functions.
7. So far, fast scoring functions cover only part of the whole receptor-ligand binding process. A more detailed picture could be obtained by taking into account properties of the unbound ligand, that is, solvation effects and energetic differences between the low-energy solution conformations and the bound conformation.

7 ACKNOWLEDGMENTS

The authors have benefited from numerous' discussions with many researchers active in the field of docking and scoring, especially Holger Gohlke (University of Marburg/ Scripps Research Institute), Ingo Muegge (Bayer), and Matthias Rarey (GMD St. Augustin).

REFERENCES

1. H. J. Boehm and G. Klebe, ***Angew. Chem. Int. Ed. Engl.,*** 35, 2588 (1996).
2. R. E. Babine and S. L. Bender, ***Chem. Rev.,*** 97, 1359 (1997).
3. J. Greer, J. W. Erickson, J. J. Baldwin, and M. D. Varney, J. ***Med. Chem.,*** 37,1035 (1994).
4. 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 (1997).

W. Lew and U. Choung, Curr. Med. Chem., **7, 663** (2000).

M. *von Itzstein, W.-Y. Wu, G. B. Kok, M. S. Pegg, J. C. Dyason, B. Jin, T. V. Phan, M. L. Smythe, H. F. White, S. W. Oliver, P. M.* **Colmant,** *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** (1993).

R. S. Bohacek, C. **McMartin,** *and W. C. Guida, Med. Res. Rev.,* **16, 3 (1996).**

R. **E. Hubbard,** *Cum. Opin. Biotechnol.,* **8,696 (1997).**

M. A. Murcko, P. R. **Caron,** *and P. S. Charifson, Ann. Rep. Med. Chem.,* **34,297** (1999).

G. Klebe, **J.** *Mol. Med.,* **78,269 (2000).**

D. A. Gschwend, W. Sirawaraporn, D. V. Santi, and I. D. Kuntz, Proteins, **29, 59 (1997).**

E. **K.** *Kick, D. C. Roe, A. G.* **Skillman,** *G. Liu, T. J. Ewing, Y. Sun, I. D. Kuntz, and J. A.* **Ellman,** *Chem. Biol.,* **4,297 (1997).**

P. Burkhard, **U.** *Hommel, M. Sanner, and M. D.* **Walkinshaw,** *J. Mol. Biol.,* **287, 853 (1999).**

J. *H. Toney, P. M. Fitzgerald, N.* **Grover-Sharma,** *S. H. Olson, W. J. May, J. G. Sundelof, D. E. Vanderwall, K. A.* **Cleary,** *S. K. Grant, J. K. Wu, J. W.* **Kozarich,** *D. L. Pompliano, and G. G. Hammond, Chem. Biol.,* **5, 185 (1998).**

S. Grueneberg, B. Wendt, and G. Klebe, Angew. Chem. Int. Ed. Engl., **40,389 (2001).**

H. J. Boehm, M. Boehringer, D. Bur, H. Gmuender, W. Huber, W. Klaus, D. Kostrewa, H. Kuehne, T. Luebbers, N. Meunier-Keller, and F. Mueller, J. Med. Chem., **43, 2664 (2000).**

S. *K. Burley, Nat. Struct. Biol.* **7** (Suppl.), **932 (2000).**

R. Abagyan and M. Totrov, Curr. Opin. Chem. Biol., **5, 375 (2001).**

I. Muegge and M. **Rarey** *in K. B. Lipkowitz and D. B. Boyd, Eds., Reviews in Computational Chemistry,* **Vol. 17,** *Wiley-VCH, New York,* **2001,** *p.* **1.**

G. R. Smith and M. J. Sternberg, Curr. Opin. Struct. Biol., **12, 28 (2002).**

C. J. Camacho and S. Vajda, Curr. Opin. Struct. Biol., **12, 36 (2002).**

A. *H.* **Elcock,** *D. Sept, and J. A. McCammon, J. Phys. Chem. B,* **105, 1504 (2001).**

B. K. *Shoichet and I. D. Kuntz, Chem. Biol.,* **3, 151 (1996).**

M. *J. Sternberg, H. A. Gabb, and R. M. Jackson, Curr. Opin. Struct. Biol.,* **8,250 (1998).**

A. V. **Filikov,** *V. Mohan, T. A. Vickers, R. H. Griffey, P. D. Cook, R. A. Abagyan, and T. L. James, J. Comput.-Aided Mol. Des.,* **14, 593 (2000).**

D. E. Clark, C. W. Murray, and J. Li in K. B. Lipkowitz and D. B. Boyd, Eds., Reviews in Computational Chemistry, **Vol. 11,** *Wiley-VCH, New York,* **1997,** *p.* **67.**

M. A. Murcko in K. B. Lipkowitz and D. B. Boyd, Eds., Reviews in Computational Chemistry, **Vol. 11,** *Wiley-VCH, New York,* **1997,** *p.* **1.**

G. Schneider and H. J. Boehm, Drug Discov. Today, **7, 64 (2002).**

W. P. Walters, M. T. **Stahl,** *and M. A. Murcko, Drug Discov. Today,* **3,160 (1998).**

T. Langer and R. D. Hoffmann, Cum. Pharm. Des., **7,509 (2001).**

W. F. van Gunsteren and H. J. C. Berendsen, Angew. Chem. Znt. Ed. Engl., **29,992** (1990).

B. Honig and A. Nicholls, Science, **268, 1144 (1995).**

G. *A. Jeffrey and W. Saenger, Hydrogen Bonding in Biological Structures, Springer-Verlag, Berlin,* **1991.**

G. A. Jeffrey, An Introduction to Hydrogen Bonding, Oxford University Press, New York, **1997.**

A. M. Davis and S. J. **Teague,** *Angew. Chem. Int. Ed. Engl.,* **38, 736 (1999).**

S. K. **Burley** *and G. A. Petsko, Adv. Protein Chem.,* **39,125 (1988).**

S. **Tsuzuki,** *K. Honda, T.* **Uchimaru,** *M.* **Mikami,** *and K. Tanabe, J. Am. Chem. Soc.,* **122, 11450 (2000).**

M. Brandl, M. S. Weiss, A. Jabs, J. Suhnel, and R. Hilgenfeld, J. Mol. Biol., **307,357 (2001).**

J. P. Gallivan and D. A. Dougherty, Proc. Natl. Acad. Sci. USA, **96,9459** (1999).

I. L. **Alberts,** *K.* **Nadassy,** *and S. J. Wodak, Protein Sci.,* **7, 1700 (1998).**

N. Borkakoti, Curr. Opin. Drug Disc. Dev., **2, 449 (1999).**

A. Ben-Naim, Hydrophobic Interactions, Plenum, New York, **1980.**

C. **Tanford,** *The Hydrophobic Effect, John Wiley & Sons, New York,* **1980.**

Ajay and M. A. Murcko, J. Med. Chem., **38, 4953 (1995).**

M. K. Gilson, J. A. Given, B. L. Bush, and J. A. McCammon, Biophys. J., **72, 1047** (1997).

J. **E. Ladbury** *and B. Z. Chowdhry, Chem. Biol., 3,* **791 (1996).**

47. T. Wiseman, S. Williston, J. F. Brandts, and L. N. Lin, *Anal. Biochem.*, **179, 131 (1989).**
48. S. H. Sleigh, P. R. Seavers, A. J. Wilkinson, J. E. Ladbury, and J. R. Tame, ***J. Mol. Biol.*, 291,393 (1999).**
49. M. H. Parker, D. F. Ortwine, P. M. O'Brien, E. A. Lunney, C. A. Banotai, W. T. Mueller, P. McConnell, and C. G. Brouillette, ***Bioorg. Med. Chem. Lett.*, 10,2427** (2000).
50. D. H. Williams, D. P. O'Brien, and B. Bardsley, ***J. Am. Chem. Soc.*, 123, 737 (2001).**
51. P. Gilli, V. Ferretti, G. Gilli, and P. A. Brea, ***J. Phys. Chem.*, 98, 1515 (1994).**
52. J. D. Dunitz, ***Chem. Biol.*, 2, 709 (1995).**
53. K. Sharp, ***Protein Sci.*, 10,661 (2001).**
54. P. C. Weber, J. J. Wendoloski, M. W. Pantoliano, and F. R. Salemme, ***J. Am. Chem. Soc.*, 114, 3197 (1992).**
55. F. Dullweber, M. T. Stubbs, D. Musil, J. Stuerzebecher, and G. Klebe, ***J. Mol. Biol.*, 313,593 (2001).**
56. J. R. Tame, ***J. Cornput.-Aided Mol. Des.*, 13**, 99 **(1999).**
57. A. R. Khan, J. C. Parrish, M. E. Fraser, W. W. Smith, P. A. Bartlett, and M. N. James, ***Biochemistry*, 37,16839 (1998).**
58. H. Mack, T. Pfeiffer, W. Hornberger, H. J. Boehm, and H. W. Hoeffken, ***J. Enzyme Znhib.*, 9**, 73 (1995).
59. Y. W. Chen, A. R. Fersht, and K. Henrick, ***J. Mol. Biol.*, 234, 1158 (1993).**
60. P. R. Connelly, R. A. Aldape, F. J. Bruzzese, S. P. Chambers, M. J. Fitzgibbon, M. A. Fleming, S. Itoh, D. J. Livingston, M. A. Navia, J. A. Thomson, and K. P. Wilson, ***Proc. Natl. Acad. Sci. USA*, 91, 1964 (1994).**
61. A. R. Fersht, J. P. Shi, J. Knill-Jones, D. M. Lowe, A. J. Wilkinson, D. M. Blow, P. Brick, P. Carter, M. M. Waye, and G. Winter, ***Nature*, 314,235 (1985).**
62. U. Obst, D. W. Banner, L. Weber, and F. Diederich, ***Chem. Biol.*, 4, 287 (1997).**
63. B. P. Morgan, J. M. Scholtz, M. D. Ballinger, I. D. Zipkin, and P. A. Bartlett, ***J. Am. Chem. Soc.*, 113, 297 (1991).**
64. B. A. Shirley, P. Stanssens, U. Hahn, and C. N. Pace, ***Biochemistry*, 31, 725 (1992).**
65. H. Kubinyi in B. Testa, H. van de Waterbeemd, G. Folkers, and R. Guy, Eds., ***Pharmacokinetic Optimization in Drug Research*,** Wiley-VCH, Weinheim, Germany, **2001,** p. **513.**
66. H.-J. Schneider, T. Schiestel, and P. Zimmermann, ***J. Am. Chem. Soc.*, 114,7698 (1992).**
67. A. C. Tissot, S. Vuilleumier, and A. R. Fersht, ***Biochemistry*, 35, 6786 (1996).**
68. J. D. Dunitz, ***Science*, 264,670 (1994).**
69. C. Chothia, ***Nature*, 254,304 (1975).**
70. F. M. Richards, ***Annu. Rev. Biophys. Bioeng.*, 6, 151** (1977).
71. K. A. Sharp, A. Nicholls, R. Friedman, and B. Honig, ***Biochemistry*, 30,9686 (1991).**
72. M. S. Searle and D. H. Williams, ***J. Am. Chem. Soc.*, 114, 10690 (1992).**
73. M. S. Searle, D. H. Williams, and U. Gerhard, ***J. Am. Chem. Soc.*, 114,10697 (1992).**
74. M. A. Hossain and H. J. Schneider, ***Chem. Eur. J.*,5, 1284 (1999).**
75. J. Hermans and L. Wang, ***J. Am. Chem. Soc.*, 119,2702 (1997).**
76. K. P. Murphy, D. Xie, K. S. Thompson, L. M. Amzel, and E. Freire, ***Proteins*, 18**, 63 **(1994).**
77. K. A. Dill, ***J. Biol. Chem.*, 272, 701** (1997).
78. G. Folkers, Ed., ***Pharm. Acta Helv.*, 69, 175 (1995).**
79. D. E. J. Koshland, ***Angew. Chem. Znt. Ed. Engl.*, 33,2408 (1994).**
80. M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, and R. P. Mee, ***J. Cornput.-Aided Mol. Des.*, 11,425 (1997).**
81. I. K. McDonald and J. M. Thornton, ***J. Mol. Biol.*, 238, 777 (1994).**
82. M. Totrov and R. Abagyan in R. B. Raffa, Ed., ***Drug-Receptor Thermodynamics: Introduction and Applications*,** John Wiley & Sons, Chichester, **2001,** p. **603.**
83. I. D. Kuntz, E. C. Meng, and B. K. Shiochet, ***Acc. Chem. Res.*, 27, 117 (1994).**
84. A. R. Leach and M. M. Hann, ***Drug Discov. Today*, 5, 326 (2000).**
85. R. A. Lewis, S. D. Pickett, and D. E. Clark in K. B. Lipkowitz and D. B. Boyd, Eds., ***Reviews in Computational Chemistry*, Vol. 16,** Wiley-VCH, New York, **2000,** p. **1.**
86. P. S. Charifson, J. J. Corkery, M. A. Murcko, and W. P. Walters, ***J. Med. Chem.*, 42, 5100 (1999).**
87. G. E. Terp, B. N. Johansen, I. T. Christensen, and F. S. Jorgensen, ***J. Med. Chem.*, 44, 2333 (2001).**
88. T. P. Lybrand, ***Curr. Opin. Struct. Biol.*, 5,224 (1995).**
89. G. Jones and P. Willett, ***Curr. Opin. Biotechnol.*, 6,652 (1995).**
90. T. Lengauer and M. Rarey, ***Curr. Opin. Struct. Biol.*, 6,402 (1996).**

R. Rosenfeld, S. Vajda, and C. DeLisi, *Annu. Rev. Biophys. Biomol. Struct.,* **24, 677 (1995).**

P. Bamborough and F. E. Cohen, *Curr. Opin. Struct. Biol.,* **6, 236 (1996).**

J. S. Dixon and J. M. Blaney in Y. C. Martin and P. Willett, Eds., *Designing Bioactive Molecules: Three-Dimensional Techniques and Applications,* American Chemical Society, Washington, DC, **1997,** p. **175.**

P. J. Gane and P. M. Dean, *Curr. Opin. Struct. Biol.,* **10,401 (2000).**

C. A. Sotriffer, W. Flader, R. H. Winger, B. M. Rode, K. R. Liedl, and J. M. Varga, *Methods,* **20,280 (2000).**

R. B. Russell and D. S. Eggleston, *Nat. Struct. Biol.***7** (Suppl.), **928** (2000).

M. Hendlich, F. Rippmann, and G. Barnickel, J. *Mol. Graph. Model.,* **15, 359 (1997).**

G. P. Brady, Jr. and P. F. Stouten, J. *Comput.-Aided Mol. Des.,* **14,383 (2000).**

C. A. Orengo, A. E. Todd, and J. M. Thornton, *Curr. Opin. Struct. Biol.,* **9, 374** (1999).

J. M. Thornton, A. E. Todd, D. Milburn, N. Borkakoti, and C. *A.* Orengo, *Nat. Struct. Biol.***7** (Suppl.), **991 (2000).**

S. Schmitt, M. Hendlich, and G. Klebe, *Angew. Chem. Znt. Ed. Engl.,* **40,3141 (2001).**

J. Ruppert, W. Welch, and A. N. Jain, *Protein Sci.,* **6,524 (1997).**

F. Jiang and S. H. Kim, J. *Mol. Biol.,* **219, 79 (1991).**

D. Fischer, S. L. Lin, H. L. Wolfson, and R. Nussinov, J. *Mol. Biol.,* **248,459 (1995).**

P. Burkhard, P. Taylor, and M. D. Walkinshaw, J. *Mol. Biol.,* **277,449 (1998).**

I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, and T. E. Ferrin, J. *Mol. Biol.,* **161, 269 (1982).**

C. M. Oshiro and I. D. Kuntz, *Proteins,* **30,321** (1998).

H. J. Boehm, J. *Cornput.-Aided Mol. Des.,* **6, 593 (1992).**

H. J. Boehm, J. *Cornput.-Aided Mol. Des.,* **6, 61 (1992).**

M. Rarey, B. Kramer, T. Lengauer, and G. Klebe, J. *Mol. Biol.,* **261, 470 (1996).**

V. Schnecke and L. A. Kuhn, *Perspect. Drug Discov. Des.,* **20, 171 (2000).**

D. J. Diller and K. M. Merz, Jr., *Proteins,* **43, 113 (2001).**

D. S. Goodsell and *A.* J. Olson, *Proteins,* **8,195 (1990).**

114. G. M. Morris, D. S. Goodsell, R. Huey, and A. J. Olson, J. *Cornput.-Aided Mol. Des.,* **10, 293 (1996).**

115. G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, and A. J. Olson, J. *Comput. Chem.,* **19,1639 (1998).**

116. R. Abagyan, M. Trotov, and D. Kuznetsov, *J. Comput. Chem.,* **15,488 (1994).**

117. M. Totrov and R. Abagyan, *Proteins* (Suppl.), **215 (1997).**

118. J. Y. Trosset and H. *A.* Scheraga, J. *Comput. Chem.,* **20,412 (1999).**

119. J. Y. Trosset and H. A. Scheraga, *Proc. Natl. Acad. Sci. USA,* **95,8011 (1998).**

120. F. S. Kuhl, G. M. Crippen, and W. G. Richards, J. *Comput. Chem.,* **5,** 24 **(1984).**

121. M. D. Miller, S. K. Kearsley, D. J. Underwood, and R. P. Sheridan, J. *Cornput.-Aided Mol. Des.,* **8, 153 (1994).**

122. D. M. Lorber and B. K. Shoichet, *Protein Sci.,* **7,938 (1998).**

123. M. McGann (2001). FRED [Online]. OpenEye. http://www.eyesopen.com/fred.html [2001, Sept 25].

124. Y. P. Pang and A. P. Kozikowski, J *Comput.-Aided Mol. Des.,* **8,669 (1994).**

125. Y. P. Pang, E. Perola, K. Xu, and F. G. Prendergast, J. *Comput. Chem.,* **22, 1750 (2001).**

126. E. C. Meng, D. A. Gschwend, J. M. Blaney, and I. D. Kuntz, *Proteins,* **17,266 (1993).**

127. D. A. Gschwend and I. D. Kuntz, *J. Comput.-Aided Mol. Des.,* **10, 123 (1996).**

128. R. L. DesJarlais, R. P. Sheridan, J. S. Dixon, I. D. Kuntz, and R. Venkataraghavan, *J. Med. Chem.,* **29,2149** (1986).

129. J. Wang, P. A. Kollman, and I. D. Kuntz, *Proteins,* **36, 1 (1999).**

130. M. Rarey, B. Kramer, and T. Lengauer, J. *Cornput.-Aided Mol. Des.,* **11,369 (1997).**

131. T. J. Ewing, S. Makino, A. G. Skillman, and I. D. Kuntz, J. *Cornput.-Aided Mol. Des.,* **15, 411 (2001).**

132. R. L. DesJarlais, R. P. Sheridan, G. L. Seibel, J. S. Dixon, I. D. Kuntz, and R. Venkataraghavan, J. *Med. Chem.,* **31,722 (1988).**

133. D. J. Bacon and J. Moult, J. *Mol. Biol.,* **225, 849 (1992).**

134. A. Wallqvist and D. G. Covell, *Proteins,* **25,403** (1996).

135. M. Y. Mizutani, N. Tomioka, and A. Itai, *J. Mol. Biol.,* **243, 310 (1994).**

136. B. B. Goldman and W. T. Wipke, ***Proteins,*** 38, 79 (2000).
137. S. Makino and I. D. Kuntz, ***J. Comput. Chem.,*** 18, 1812 (1997).
138. M. Rarey, S. Wefing, and T. Lengauer, ***J. Comput.-Aided Mol. Des.,*** **10**, 41 (1996).
139. B. Kramer, M. Rarey, and T. Lengauer, ***Proteins,*** 37, 228 (1999).
140. G. Klebe and T. Mietzner, ***J. Cornput.-Aided Mol. Des.,*** 8,583 (1994).
141. G. Klebe, ***J. Mol. Biol.,*** 237,212 (1994).
142. R. L. DesJarlais and J. S. Dixon, ***J. Comput.-Aided Mol. Des.,*** 8,231 (1994).
143. B. K. Shoichet and I. D. Kuntz, ***Protein Eng.,*** 6, 723 (1993).
144. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, ***J. Chem. Phys.,*** 21, 1087 (1953).
145. T. N. Hart, S. R. Ness, and R. J. Read, ***Proteins*** (Suppl.), 205 (1997).
146. T. N. Hart and R. J. Read, ***Proteins,*** 13, 206 (1992).
147. M. Liu and S. Wang, ***J. Cornput.-Aided Mol. Des.,*** 13,435 (1999).
148. Z. Li and H. A. Scheraga, ***Proc. Natl. Acad. Sci. USA,*** 84,6611 (1987).
149. R. Abagyan and P. Argos, ***J. Mol. Biol.,*** 225, 519 (1992).
150. A. Caflisch, S. Fischer, and M. Karplus, ***J. Comput. Chem.,*** 18, 723 (1997).
151. I. Apostolakis, A. Plueckthun, and A. Caflisch, ***J. Comput. Chem.,*** **19**, 21 (1998).
152. C. McMartin and R. S. Bohacek, ***J. Comput.-Aided Mol. Des.,*** 11,333 (1997).
153. M. Karplus and J. A. McCammon, ***Annu. Rev. Biochem.,*** **52**, 263 (1983).
154. W. F. van Gunsteren and A. E. Mark, ***Eur. J. Biochem.,*** 204,947 (1992).
155. W. F. van Gunsteren, P. H. Huenenberger, A. E. Mark, P. E. Smith, and I. G. Tironi, ***Comput. Phys. Commun.,*** 91,305 (1995).
156. D. Rognan, ***Perspect. Drug Discov. Des.,*** 11, 181 (1998).
157. M. E. Tuckerman and G. J. Martyna, ***J. Phys. Chem. B,*** 104, 159 (2000).
158. D. L. Beveridge and F. M. DiCapua, ***Annu. Rev. Biophys. Biophys. Chem.,*** 18,431 (1989).
159. D. A. Pearlman and P. A. Kollman in W. F. van Gunsteren and P. K. Weiner, Eds., ***Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications,*** **Vol.** 1, ESCOM, Leiden, The Netherlands, 1989, p. 101.
160. T. P. Straatsma and J. A. McCammon, ***Annu. Rev. Phys. Chem.,*** 43,407 (1992).
161. P. A. Kollman, ***Chem. Rev.,*** 93,2395 (1993).
162. T. P. Straatsma in K. B. Lipkowitz and D. B. Boyd, Eds., ***Reviews in Computational Chemistry,*** **Vol.** 9, VCH, New York, 1996, p. 81.
163. M. R. Reddy, M. D. Erion, and A. Agarwal in K. B. Lipkowitz and D. B. Boyd, Eds., ***Reviews in Computational Chemistry,*** **Vol.** ***16,*** Wiley-VCH, New York, 2000, p. 217.
164. A. Di Nola, D. Roccatano, and H. J. Berendsen, ***Proteins,*** 19, 174 (1994).
165. M. Mangoni, D. Roccatano, and A. Di Nola, ***Proteins,*** 35, 153 (1999).
166. M. Vieth, J. D. Hirst, B. N. Dominy, H. Daigler, and C. L. Brooks, ***J. Comput. Chem.,*** 19, 1623 (1998).
167. Y. Pak and S. Wang, ***J. Phys. Chem. B,*** 104,354 (2000).
168. Y. Pak, I. J. Enyedy, J. Varady, J. W. Kung, P. S. Lorenzo, P. M. Blumberg, and S. Wang, ***J. Med. Chem.,*** 44, 1690 (2001).
169. B. A. Luty, Z. R. Wasserman, P. W. F. Stouten, C. N. Hodge, M. Zacharias, and J. A. McCammon, ***J. Comput. Chem.,*** 16,454 (1995).
170. A. Miranker and M. Karplus, ***Proteins,*** **11,** 29 (1991).
171. A. Caflisch, A. Miranker, and M. Karplus, ***J. Med. Chem.,*** 36,2142 (1993).
172. R. Judson in K. B. Lipkowitz and D. B. Boyd, Eds., ***Reviews in Computational Chemistry,*** **Vol.** 10, VCH, New York, 1997.
173. D. E. Goldberg, ***Genetic Algorithms in Search, Optimization, and Machine Learning,*** Addison-Wesley, Reading, MA, 1989.
174. L. Davis, ***Handbook of Genetic Algorithms,*** Van Nostrand Reinhold, New York, 1991.
175. D. E. Clark and D. R. Westhead, ***J. Comput.-Aided Mol. Des.,*** 10, 337 (1996).
176. G. Jones, P. Willett, and R. C. Glen, ***J. Mol. Biol.,*** **245**, 43 (1995).
177. G. Jones, P. Willett, R. C. Glen, A. R. Leach, and R. Taylor, ***J. Mol. Biol.,*** 267, 727 (1997).
178. J. S. Taylor and R. M. Burnett, ***Proteins,*** 41, 173 (2000).
179. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, ***J. Comput. Chem.,*** 4, 187 (1983).
180. R. S. Judson, E. P. Jaeger, and A. M. Treasurywala, ***Theochem. J. Mol. Struct.,*** 114, 191 (1994).
181. K. P. Clark and Ajay, ***J. Comput. Chem.,*** 16, 1210 (1995).

C. M. Oshiro, I. D. Kuntz, and J. S. Dixon, *J. Cornput.-Aided Mol. Des.*, 9, 113 (1995).

M. Thormann and M. Pons, *J. Comput. Chem.*, 22, 1971 (2001).

G. Jones in P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer, 3rd, and R. P. Schreiner, Eds., *Encyclopedia of Computational Chemistry*, John Wiley & Sons, New York, 1998.

D. K. Gehlhaar, G. M. Verkhivker, P. A. Rejto, C. J. Sherman, D. B. Fogel, L. J. Fogel, and S. T. Freer, *Chem. Biol.*, 2, 317(1995).

D. R. Westhead, D. E. Clark, and C. W. Murray, *J. Cornput.-Aided Mol. Des.*, 11, 209(1997).

J. M. Yang and C. Y. Kao, *J. Comput. Chem.*, 21, 988(2000).

C. A. Baxter, C. W. Murray, D. E. Clark, D. R. Westhead, and M. D. Eldridge, *Proteins*, 33, 367 (1998).

C. A. Baxter, C. W. Murray, B. Waszkowycz, J. Li, R. A. Sykes, R. G. Bone, T. D. Perkins, and W. Wylie, *J. Chem. Inf. Comput. Sci.*, 40, 254 (2000).

L. David, R. Luo, and M. K. Gilson, *J. Comput.-Aided Mol. Des.*, 15, 157 (2001).

J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, London, 1987.

H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, *Science*, 254, 1598 (1991).

E. Freire, *Adv. Protein Chem.*, 51, 255 (1998).

B. Ma, S. Kumar, C. J. Tsai, and R. Nussinov, *Protein Eng.*, 12, 713 (1999).

M. Gerstein and W. Krebs, *Nucleic Acids Res.*, 26, 4280(1998).

H. A. Carlson and J. A. McCammon, *Mol. Pharmacol.*, 57, 213(2000).

H. Claussen, C. Buning, M. Rarey, and T. Lengauer, *J. Mol. Biol.*, 308, 377(2001).

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

A. R. Leach, *J. Mol. Biol.*, 235, 345(1994).

L. Schaffer and G. M. Verkhivker, *Proteins*, 33, 295 (1998).

M. Schapira, B. M. Raaka, H. H. Samuels, and R. Abagyan, *Proc. Natl. Acad. Sci. USA*, 97, 1008 (2000).

V. Schnecke, C. A. Swanson, E. D. Getzoff, J. A. Tainer, and L. A. Kuhn, *Proteins*, 33, 74 (1998).

A. C. Anderson, R. H. O'Neil, T. S. Surti, and R. M. Stroud, *Chem. Biol.*, 8, 445(2001).

204. H. A. Carlson, K. M. Masukawa, K. Rubins, F. D. Bushman, W. L. Jorgensen, R. D. Lins, J. M. Briggs, and J. A. McCammon, *J. Med. Chem.*, 43, 2100(2000).
205. H. A. Carlson, K. M. Masukawa, and J. A. McCammon, *J. Phys. Chem. A*, 103, 10213(2000).
206. D. Bouzida, P. A. Rejto, S. Arthurs, A. B. Colson, S. T. Freer, D. K. Gehlhaar, V. Larson, B. A. Luty, P. W. Rose, and G. M. Verkhivker, *Int. J. Quantum Chem.*, 72, 73 (1999).
207. R. M. Knegtel, I. D. Kuntz, and C. M. Oshiro, *J. Mol. Biol.*, 266, 424(1997).
208. F. Osterberg, G. M. Morris, M. F. Sanner, A. J. Olson, and D. S. Goodsell, *Proteins*, 46, 34 (2002).
209. H. B. Broughton, *J. Mol. Graph. Model.*, 18, 247 (2000).
210. I. Kolossváry and W. C. Guida, *J. Am. Chem. Soc.*, 118, 5011(1996).
211. I. Kolossváry and W. C. Guida, *J. Comput. Chem.*, 20, 1671 (1999).
212. J. Israelachvili and H. Wennerstrom, *Nature*, 379, 219(1996).
213. M. Levitt and B. H. Park, *Structure*, 1, 223 (1993).
214. J. E. Ladbury, *Chem. Biol.*, 3, 973(1996).
215. P. Y. Lam, P. K. Jadhav, C. J. Eyermann, C. N. Hodge, Y. Ru, L. T. Bacheler, J. L. Meek, M. J. Otto, M. M. Rayner, Y. N. Wong, C. H. Chang, P. C. Weber, D. A. Jackson, T. R. Sharpe, and S. Ericksonviitanen, *Science*, 263, 380 (1994).
216. M. Rarey, B. Kramer, and T. Lengauer, *Proteins*, 34, 17 (1999).
217. M. L. Raymer, P. C. Sanschagrin, W. F. Punch, S. Venkataraman, E. D. Goodman, and L. A. Kuhn, *J. Mol. Biol.*, 265, 445(1997).
218. P. J. Goodford, *J. Med. Chem.*, 28, 849(1985).
219. R. C. Wade and P. J. Goodford, *J. Med. Chem.*, 36, 148 (1993).
220. R. C. Wade, K. J. Clark, and P. J. Goodford, *J. Med. Chem.*, 36, 140 (1993).
221. W. E. Minke, D. J. Diller, W. G. Hol, and C. L. Verlinde, *J. Med. Chem.*, 42, 1778 (1999).
222. M. S. Rao and A. J. Olson, *Proteins*, 34, 173 (1999).
223. P. Pospisil, L. Scapozza, and G. Folkers in H. D. Hoeltje and W. Sippl, Eds., *Rational Approaches to Drug Design. Proceedings of the 13th European Symposium on Quantitative Structure-Activity Relationships*, Prous Science, Barcelona/Philadelphia, 2001, p. 92.
224. R. M. Knegtel, D. M. Bayada, R. A. Engh, W. von der Saal, V. J. van Geerestein, and P. D. Grootenhuis, *J. Comput.-Aided Mol. Des.*, 13, 167 (1999).

225. *P. Tao and L.* Lai, *J. Cornput.-Aided Mol. Des.*, **15, 429 (2001).**

226. *H. Gohlke, M. Hendlich, and G. Klebe, J. Mol. Biol.*, **295, 337 (2000).**

227. *O.* Roche, *R. Kiyama, and C. L. Brooks,* 3rd, *J. Med. Chem.*, **44,3592 (2001).**

228. *C. A. Sotriffer, H. H. Ni, and J. A.* McCammon, *J.* Am. *Chem. Soc.*, **122,6136 (2000).**

229. *S. Ha, R. Andreani, A. Robbins, and I. Muegge, J. Cornput.-Aided Mol. Des.*, **14,435 (2000).**

230. *C. Bissantz, G. Folkers, and D. Rognan, J. Med. Chem.*, **43,4759 (2000).**

231. *J. S. Dixon, Proteins (Suppl.),* **198 (1997).**

232. *N. C. Strynadka, M. Eisenstein, E.* Katchalski-*Katzir, B. K. Shoichet, I. D. Kuntz, R.* Abagyan, *M. Totrov, J. Janin, J. Cherfils, F.* Zimmerman, *A. Olson, B. Duncan, M. Rao, R. Jackson, M. Sternberg, and M. N. James, Nat. Struct. Biol.*, **3,233 (1996).**

233. **B.** *Kramer, M.* Rarey, *and T. Lengauer, Proteins (Suppl.),* **221 (1997).**

234. *V. Sobolev, T. M. Moallem, R. C. Wade, G. Vriend, and M. Edelman, Proteins (Suppl.),* **210 (1997).**

235. *H. Kubinyi, Ed., 3D QSAR in Drug Design. Theory, Methods, and Applications, ESCOM,* Leiden, *The Netherlands,* **1993.**

236. *H. Kubinyi, G. Folkers,* and Y. *C. Martin, Eds., 3D QSAR in Drug Design. Recent Advances,* Vol. 3, Kluwer/ESCOM, *Dordrecht, The Netherlands,* **1998.**

237. *H. Kubinyi, QSAR, Hansch Analysis and Related Approaches, VCH, Weinheim, Germany,* **1993.**

238. *W. Sippl,* **J.** *Cornput.-Aided Mol. Des.*, **14, 559 (2000).**

239. *C. L. Waller, T. I. Oprea, A. Giolitti, and G. R. Marshall, J. Med. Chem.*, **36,4152 (1993).**

240. *A. M. Gamper, R. H. Winger, K. R. Liedl, C. A. Sotriffer, J. M. Varga, R. T.* Kroemer, *and B. M. Rode, J. Med. Chem.*, **39,3882 (1996).**

241. *R. C. Wade in H. D. Hoeltje and W. Sippl, Eds., Rational Approaches to Drug Design. Proceedings of the 13th European Symposium on Quantitative Structure-Activity Relationships,* Prous *Science,* Barcelona/Philadelphia, **2001,** *p.* **23.**

242. *A. R. Ortiz, M. T. Pisabarro, F. Gago, and R. C. Wade, J. Med. Chem.*, **38,2681 (1995).**

243. *R. C. Wade, A. R. Ortiz, and F. Gago, Perspect. Drug Discov. Des.*, **9, 19 (1998).**

244. *T. Wang and R. C. Wade in H. D. Hoeltje and W. Sippl, Eds., Rational Approaches to Drug Design. Proceedings of the 13th European Symposium on Quantitative Structure-Activity Relationships,* Prous *Science,* Barcelona/Philadelphia, **2001,** *p.* **78.**

245. *G. Cruciani and K. A. Watson, J. Med. Chem.*, **37,2589 (1994).**

246. *W. Sippl and H. D. Hoeltje, J. Mol. Struct.-Theochem.*, **503,** 31 **(2000).**

247. *M. Vieth and D. J. Cummins, J. Med. Chem.*, **43, 3020 (2000).**

248. *M.* Wojciechowski *and J. Skolnick, J. Comput. Chem.*, **23, 189 (2002).**

249. *S. Moro, A. H. Li, and K. A. Jacobson, J. Chem. Inf. Comput. Sci.*, **38, 1239 (1998).**

250. *S.* Moro, *D. Guo, E. Camaioni, J. L. Boyer, T. K. Harden, and K. A. Jacobson, J. Med. Chem.*, **41, 1456 (1998).**

251. *R. Kiyama, Y. Tamura, F. Watanabe, H.* Tsuzuki, *M. Ohtani, and M.* Yodo, *J. Med. Chem.*, **42, 1723 (1999).**

252. *C. A. Sotriffer, W. Flader, A. Cooper, B. M. Rode, D. S. Linthicum, K. R. Liedl, and J. M. Varga, Biophys. J.*, **76,2966** (1999).

253. *A. Schafferhans and G. Klebe, J. Mol. Biol.*, **307,407 (2001).**

254. *S. Kearsley and G. Smith, Tetrahedron Comput. Methodol.*, **3, 615 (1990).**

255. *G. Klebe, T. Mietzner, and F. Weber, J. Comput.-Aided Mol. Des.*, **8, 751 (1994).**

256. *G. Klebe, T. Mietzner, and F. Weber, J. Comput.-Aided Mol. Des.*, **13,** 35 **(1999).**

257. *P. A.* Kollman, *Acc. Chem. Res.*, **29,461 (1996).**

258. *M. L. Lamb and W. L. Jorgensen, Curr. Opin. Chem. Biol.*, **1,449 (1997).**

259. *M. K. Gilson, J. A. Given, and M. S. Head, Chem. Biol.*, **4, 87 (1997).**

260. *J. Aqvist, C. Medina, and J. E. Samuelsson, Protein Eng.*, **7, 385 (1994).**

261. *T. Hansson, J. Marelius, and J. Aqvist, J. Comput.-Aided Mol. Des.*, **12, 27 (1998).**

262. *J. Aqvist, V. B. Luzhkov, and B.* O. *Brandsdal, Acc. Chem. Res.*, **35,358 (2002).**

263. **R.** *C. Rizzo, J. Tirado-Rives, and W. L.* Jorgensen, *J. Med. Chem.*, **44, 145 (2001).**

264. *A. E. Mark and W. F. van Gunsteren, J. Mol. Biol.*, **240, 167 (1994).**

265. *D. Williams and B. Bardsley, Perspect. Drug Discov. Des.*, **17, 43 (1999).**

266. *P. R.* Andrews, *D. J. Craik, and J. L. Martin, J. Med. Chem.*, **27, 1648 (1984).**

267. *H. J. Schneider, Chem. Soc. Rev.*, **23, 227 (1994).**

268. T. J. Stout, C. R. Sage, and R. M. Stroud, Structure, 6, 839 (1998).

269. 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, and S. J. deSolms, J. Med. Chem., 38,305 (1995).

270. P. D. J. Grootenhuis and P. J. M. van Galen, Acta Cryst. D, 51,560 (1995).

271. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner, J. Am. Chem. Soc., 106, 765 (1984).

272. S. J. Weiner, P. A. Kollman, D. T. Nguyen, and D. A. Case, J. Comput. Chem., 7,230 (1986).

273. E. C. Meng, B. K. Shoichet, and I. D. Kuntz, J. Comput. Chem., 13,505 (1992).

274. M. Vieth, J. D. Hirst, A. Kolinski, and C. L. Brooks, J. Comput. Chem., 19,1612 (1998).

275. N. Majeux, M. Scarsi, J. Apostolakis, C. Ehrhardt, and A. Caflisch, Proteins, 37, 88

276. B. K. Shoichet, A. R. Leach, and I. D. Kuntz, Proteins, 34, 4 (1999).

277. W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson, J. Am. Chem. Soc., 112, 6127

278. X. Zou, Y. Sun, and I. D. Kuntz, J. Am. Chem. Soc., 121,8033 (1999).

279. M. L. Lamb, K. W. Burdick, S. Toba, M. M. Young, A. G. Skillman, X. Zou, J. R. Arnold, and I. D. Kuntz, Proteins, 42,296 (2001).

280. M. Schapira, M. Totrov, and R. Abagyan, J. Mol. Recognit., 12, 177 (1999).

281. T. Zhang and D. E. Koshland, Jr., Protein Sci., 5, 348 (1996).

282. P. H. Hunenberger, V. Helms, N. Narayana, S. S. Taylor, and J. A. McCammon, Biochemistry, 38,2358 (1999).

283. D. A. Pearlman and P. S. Charifson, J. Med. Chem., 44, 502 (2001).

284. D. A. Pearlman, J. Med. Chem., 42, 4313

285. J. Bostrom, P. O. Norrby, and T. Liljefors, J. Cornput.-Aided Mol. Des., 12,383 (1998).

286. M. Vieth, J. D. Hirst, and C. L. Brooks, 3rd, J. Cornput.-Aided Mol. Des., 12, 563 (1998).

287. M. Stahl and M. Rarey, J. Med. Chem., 44, 1035 (2001).

288. I. D. Kuntz, K. Chen, K. A. Sharp, and P. A. Kollman, Proc. Natl. Acad. Sci. USA, 96,9997

289. J. D. Hirst, Curr. Opin. Drug Disc. Dev., 1, 28

290. R. M. A. Knegtel and P. D. J. Grootenhuis, Perspect. Drug Discov. Des., 9–11, 99 (1998).

291. T. I. Oprea and G. R. Marshall, Perspect. Drug Discov. Des., 9–11, 3 (1998).

292. H. J. Boehm and M. Stahl, Med. Chem. Res., 9, 445 (1999).

293. R. Wang, L. Liu, L. Lai, and Y. Tang, J. Mol. Model., 4, 379 (1998).

294. H. J. Boehm, J. Cornput.-Aided Mol. Des., 8, 243 (1994).

295. C. W. Murray, T. R. Auton, and M. D. Eldridge, J. Cornput.-Aided Mol. Des., 12,503 (1998).

296. H. J. Boehm, J. Cornput.-Aided Mol. Des., 12, 309 (1998).

297. A. N. Jain, J. Cornput.-Aided Mol. Des., 10,427 (1996).

298. D. N. Boobbyer, P. J. Goodford, P. M. McWhinnie, and R. C. Wade, J. Med. Chem., 32, 1083 (1989).

299. M. Stahl, Perspect. Drug Discov. Des., 20, 83 (2000).

300. H. J. Boehm, D. W. Banner, and L. Weber, J. Conput.-Aided Mol. Des., 13, 51 (1999).

301. M. Matsumara, W. J. Becktel, and B. W. Matthews, Nature, 334,406 (1988).

302. V. Nauchitel, M. C. Villaverde, and F. Sussman, Protein Sci., 4, 1356 (1995).

303. L. Wesson and D. Eisenberg, Protein Sci., 1, 227 (1992).

304. S. Vajda, Z. Weng, R. Rosenfeld, and C. DeLisi, Biochemistry, 33, 13977 (1994).

305. C. Zhang, G. Vasmatzis, J. L. Cornette, and C. DeLisi, J. Mol. Biol., 267, 707 (1997).

306. S. Miyazawa and R. L. Jernigan, Macromolecules, 18,534 (1985).

307. R. D. Head, M. L. Smythe, T. I. Oprea, C. L. Waller, S. M. Green, and G. R. Marshall, J. Am. Chem. Soc., 118,3959 (1996).

308. E. C. Meng, I. D. Kuntz, D. J. Abraham, and G. E. Kellogg, J. Cornput.-Aided Mol. Des., 8, 299 (1994).

309. H. Gohlke and G. Klebe, Curr. Opin. Struct. Biol., 11,231 (2001).

310. M. J. Sippl, J. Comput.-Aided Mol. Des., 7,473 (1993).

311. G. Verkhivker, K. Appelt, S. T. Freer, and J. E. Villafranca, Protein Eng., 8,677 (1995).

312. A. Wallqvist, R. L. Jernigan, and D. G. Covell, Protein Sci., 4, 1881 (1995).

313. R. S. DeWitte and E. I. Shakhnovich, J. Am. Chem. Soc., 118,11733 (1996).

314. R. S. DeWitte, A. V. Ishchenko, and E. I. Shakhnovich, J. Am. Chem. Soc., 119, 4608 (1997).

315. J. B. O. Mitchell, R. A. Laskowski, A. Alex, and J. M. Thornton, J. Comput. Chem., 20, 1165 (1999).

316. J. B. O. Mitchell, R. A. Laskowski, A. Alex, M. J. Forster, and J. M. Thornton, J. Comput. Chem., 20, 1177 (1999).

317. I. Muegge and Y. C. Martin, J. Med. Chem., 42, 791 (1999).

318. H. Gohlke, M. Hendlich, and G. Klebe, Perspect. Drug Discov. Des., 20, 115 (2000).

319. I. Muegge, J. Comput. Chem., 22, 418 (2001).

320. I. Muegge, Perspect. Drug Discov. Des., 20, 99 (2000).

321. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, Nucleic Acids Res., 28, 235 (2000).

322. M. Hendlich, Acta Crystallogr. D Biol. Crystallogr., 54, 1178 (1998).

323. A. Bergner, J. Guenther, M. Hendlich, G. Klebe, and M. Verdonk, Biopolymers (Nucl. Acid Sci.), 61, 299 (2002).

324. Y. Sun, T. J. Ewing, A. G. Skillman, and I. D. Kuntz, J. Cornput.-Aided Mol. Des., 12, 597 (1998).

325. E. J. Martin, R. E. Critchlow, D. C. Spellmeyer, S. Rosenberg, K. L. Spear, and J. M. Blaney, Pharmacochem. Libr., 29, 133 (1998).

326. B. K. Shoichet, D. L. Bodian, and I. D. Kuntz, J. Comput. Chem., 13, 380 (1992).

327. T. J. A. Ewing and I. D. Kuntz, J. Comput. Chem., 18, 1175 (1997).

328. T. Ewing, Ed., DOCK User Manual, Version 4.0, Regents of the University of California, San Francisco, 1998.

329. M. Stahl and H. J. Boehm, J. Mol. Graph. Model., 16, 121 (1998).

330. C. Lemmen, T. Lengauer, and G. Klebe, J. Med. Chem., 41, 4502 (1998).

331. C. Lemmen, A. Zien, R. Zimmer, and T. Lengauer, Pac. Symp. Biocomput., 482 (1999).

332. G. R. Desiraju and T. Steiner, The Weak Hydrogen Bond in Chemistry and Biology, Oxford University Press, Oxford, 1999.

333. G. Parkinson, A. Gunasekera, J. Vojtechovsky, X. Zhang, T. A. Kunkel, H. Berman, and R. H. Ebright, Nat. Struct. Biol., 3, 837 (1996).

334. J. P. Gallivan and D. A. Dougherty, J. Am. Chem. Soc., 122, 870 (2000).

335. G. B. McGaughey, M. Gagne, and A. K. Rappe, J. Biol. Chem., 273, 15458 (1998).

336. C. Chipot, R. Jaffe, B. Maigret, D. A. Pearlman, and P. A. Kollman, J. Am. Chem. Soc., 118, 11217 (1996).

337. T. G. Davies, R. E. Hubbard, and J. R. Tame, Protein Sci., 8, 1432 (1999).

338. G. Klebe, F. Dullweber, and H. J. Boehm in R. B. Raffa, Ed., Drug-Receptor Thermodynamics: Introduction and Applications, John Wiley & Sons, Chichester, 2001, p. 83.

339. I. Muegge, Med. Chem. Res., 9, 490 (1999).

340. I. Nobeli, J. B. O. Mitchell, A. Alex, and J. M. Thornton, J. Comput. Chem., 22, 673 (2001).

341. I. Massova and P. A. Kollman, Perspect. Drug Discov. Des., 18, 113 (2000).

342. P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham, 3rd, Acc. Chem. Res., 33, 889 (2000).

343. I. Massova and P. A. Kollman, J. Am. Chem. Soc., 121, 8133 (1999).

344. O. A. Donini and P. A. Kollman, J. Med. Chem., 43, 4180 (2000).

345. B. Kuhn and P. A. Kollman, J. Med. Chem., 43, 3786 (2000).

346. B. Kuhn and P. A. Kollman, J. Am. Chem. Soc., 122, 3909 (2000).

347. N. Froloff, A. Windemuth, and B. Honig, Protein Sci., 6, 1293 (1997).

348. J. Shen, J. Med. Chem., 40, 2953 (1997).

349. C. J. Woods, M. A. King, and J. W. Essex, J. Cornput.-Aided Mol. Des., 15, 129 (2001).

350. G. Archontis, T. Simonson, and M. Karplus, J. Mol. Biol., 306, 307 (2001).

351. M. L. Verdonk, J. C. Cole, and R. Taylor, J. Mol. Biol., 289, 1093 (1999).

352. M. L. Verdonk, J. C. Cole, P. Watson, V. Gillet, and P. Willett, J. Mol. Biol., 307, 841 (2001).

353. D. E. Clark and S. D. Pickett, Drug Discov. Today, 5, 49 (2000).

354. I. Muegge, S. L. Heald, and D. Brittelli, J. Med. Chem., 44, 1841 (2001).

355. Unity Chemical Information Software, Tripos Inc., St. Louis, MO.

356. Y. Kurogi and O. F. Guener, Curr. Med. Chem., 8, 1035 (2001).

357. M. Rarey and J. S. Dixon, J. Cornput.-Aided Mol. Des., 12, 471 (1998).

358. M. Rarey and M. Stahl, J. Cornput.-Aided Mol. Des., 15, 497 (2001).

A. I. Su, D. M. Lorber, G. S. Weston, W. A. Baase, B. W. Matthews, and B. K. Shoichet, *Proteins*, **42, 279 (2001).**

Protherics **(2000).** DockCrunch [Online]. Protherics Molecular Design Ltd. http://www.protherics.com/crunch/ [2001, Sept **25].**

H. J. Boehm and M. Stahl, *Curr. Opin. Chem. Biol.*, **4, 283 (2000).**

C. W. Murray, D. E. Clark, T. R. Auton, M. A. Firth, J. Li, R. A. Sykes, B. Waszkowycz, D. R. Westhead, and S. C. Young, *J. Cornput.-Aided Mol. Des.*, **11, 193** (1997).

M. Rarey and T. Lengauer, *Perspect. Drug Discov. Des.*, **20, 63** (2000).

D. C. Roe and I. D. Kuntz, *J. Cornput.-Aided Mol. Des.*, **9, 269** (1995).

S. Makino, T. J. Ewing, and I. D. Kuntz, *J. Cornput.-Aided Mol. Des.*, **13, 513** (1999).

M. D. Miller, R. P. Sheridan, S. K. Kearsley, and D. J. Underwood, *Methods Enzymol.*, **241, 354 (1994).**

G. M. Verkhivker, D. Bouzida, D. K. Gehlhaar, P. A. Rejto, S. Arthurs, A. B. Colson, S. T. Freer, V. Larson, B. A. Luty, T. Marrone, and P. W. Rose, *J. Cornput.-Aided Mol. Des.*, **14, 731 (2000).**

S. **B.** Shuker, P. J. Hajduk, R. P. Meadows, and S. W. Fesik, *Science*, **274, 1531 (1996).**

I. Muegge, Y. C. Martin, P. J. Hajduk, and S. W. Fesik, *J. Med. Chem.*, **42, 2498 (1999).**

R. L. DesJarlais, G. L. Seibel, I. D. Kuntz, P. S. Furth, J. C. Alvarez, P. R. Ortiz de Montellano, D. L. DeCamp, L. M. Babe, and C. S. Craik, *Proc. Natl. Acad. Sci. USA*, **87, 6644 (1990).**

B. K. Shoichet, R. M. Stroud, D. V. Santi, I. D. Kuntz, and K. M. Perry, *Science*, **259, 1445** (1993).

L. **R.** Hoffman, I. D. Kuntz, and J. M. White, *J. Virol.*, **71, 8808** (1997).

I. Massova, P. Martin, A. Bulychev, R. Kocz, M. Doyle, B. F. Edwards, and S. Mobashery, *Bioorg. Med. Chem. Lett.*, **8, 2463** (1998).

D. Tondi, U. Slomczynska, M. P. Costi, D. M. Watterson, S. Ghelli, and B. K. Shoichet, *Chem. Biol.*, **6, 319 (1999).**

T. Toyoda, R. K. Brobey, G. Sano, T. Horii, N. Tomioka, and A. Itai, *Biochem. Biophys. Res. Commun.*, **235, 515** (1997).

Available Chemicals Directory, MDL Information Systems Inc., San Leandro, CA.

F. Abbate, C. T. Supuran, A. Scozzafava, P. Orioli, M. T. Stubbs, and G. Klebe, *J. Med. Chem.*, **45,** in press (2002).

378. Maybridge Database (1999), Maybridge Chemical Co. Ltd., UK.

379. LeadQuest Chemical Compound Libraries (2000), Tripos Inc., St. Louis, MO.

380. J. Sadowski, C. Rudolph, and J. Gasteiger, *Tetrahedron Comput. Methodol.*, **3, 537 (1990).**

381. A. Maxwell, *Biochem. Soc. Trans.*, **27, 48 (1999).**

382. D. Bailey and D. Brown, *Drug Discov. Today*, **6, 57 (2001).**

383. R. M. Eglen, G. Schneider, and H. J. Boehm in G. Schneider and H. J. Boehm, Eds., *Virtual Screening for Bioactive Molecules*, VCH, Weinheim, Germany, **2000,** p. **1.**

384. R. Lahana, *Drug Discov. Today*, **4, 447 (1999).**

385. C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, *Adv. Drug Delivery Rev.*, **23, 3 (1997).**

386. Y. Iwata, M. Arisawa, R. Hamada, Y. Kita, M. Y. Mizutani, N. Tomioka, A. Itai, and S. Miyamoto, *J. Med. Chem.*, **44, 1718 (2001).**

387. V. Sobolev, R. C. Wade, G. Vriend, and M. Edelman, *Proteins*, **25, 120** (1996).

388. W. Welch, J. Ruppert, and A. N. Jain, *Chem. Biol.*, **3, 449** (1996).

389. E. Perola, K. Xu, T. M. Kollmeyer, S. H. Kaufmann, F. G. Prendergast, and Y. P. Pang, *J Med Chem.*, **43, 401 (2000).**

390. D. S. Goodsell, G. M. Morris, and A. J. Olson, *J. Mol. Recognit.*, **9, 1 (1996).**

391. C. A. Sotriffer, H. Ni, and J. A. McCammon, *J. Med. Chem.*, **43, 4109** (2000).

392. *T.* Hou, J. Wang, L. Chen, and X. Xu, *Protein Eng.*, **12, 639** (1999).

393. B. N. Dominy and C. L. Brooks, 3rd, *Proteins*, **36, 318 (1999).**

394. I. D. Wall, A. R. Leach, D. W. Salt, M. G. Ford, and J. W. Essex, *J. Med. Chem.*, **42, 5142 (1999).**

395. D. A. Pearlman and P. S. Charifson, *J. Med. Chem.*, **44, 3417 (2001).**

396. S. S. So and M. Karplus, *J. Conput.-Aided Mol. Des.*, **13, 243 (1999).**

397. M. Rarey, B. Kramer, and T. Lengauer, *Bioinformatics*, **15, 243 (1999).**

398. Y. Takamatsu and A. Itai, *Proteins*, **33, 62 (1998).**

399. D. Rognan, S. L. Lauemoller, A. Holm, S. Buus, and V. Tschinke, *J. Med. Chem.*, **42, 4650 (1999).**

400. G. E. Kellogg, G. S. Joshi, and D. J. Abraham, *Med. Chem. Res.*, **1, 444** (1991).