Chemistry & Biology, Vol. 10, 787–797, September, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.chembiol.2003.09.002

The Process of Structure-Based Drug Design

Review

Amy C. Anderson*
Dartmouth College
Department of Chemistry
Burke Laboratories
Hanover, New Hampshire 03755

Summary

The field of structure-based drug design is a rapidly growing area in which many successes have occurred in recent years. The explosion of genomic, proteomic, and structural information has provided hundreds of new targets and opportunities for future drug lead discovery. This review summarizes the process of structure-based drug design and includes, primarily, the choice of a target, the evaluation of a structure of that target, the pivotal questions to consider in choosing a method for drug lead discovery, and evaluation of the drug leads. Key principles in the field of structure-based drug design will be illustrated through a case study that explores drug design for AmpC β-lactamase.

Introduction

During the early 1980s, the ability to rationally design drugs using protein structures was an unrealized goal for many structural biologists. The first projects were underway in the mid-80s, and by the early 1990s the first success stories were published [1–3]. Today, even though there is still quite a bit of fine-tuning necessary to perfect the process, structure-based drug design is an integral part of most industrial drug discovery programs [4] and is the major subject of research for many academic laboratories.

The completion of the human genome project, the start of both the proteomics and structural genomics revolutions, and developments in information technology are fueling an even greater opportunity for structure-based drug design to be part of the success story in the discovery of new drug leads. Excellent drug targets are identified at an increased pace using developments in bioinformatics. The genes for these targets can be cloned quickly, and the protein expressed and purified to homogeneity. Advances in high-throughput crystallography, such as automation at all stages, more intense synchrotron radiation, and new developments in phase determination, have shortened the timeline for determining structures. Structure determination using nuclear magnetic resonance (NMR) has also seen a number of advances in the past years, including magnet and probe improvements, automated assignment [5–7], and new experimental methods to determine larger structures [8]. Faster computers and the availability of relatively inexpensive clusters of computers have increased the speed at which drug leads can be identified and evaluated in silico.

Structure-based drug design is most powerful when it is a part of an entire drug lead discovery process. A review by J. Antel [9] states that the combination of combinatorial chemistry and structure-based design can lead to the parallel synthesis of focused compound libraries. It is also important to consider that structure-based drug design directs the discovery of a drug lead, which is not a drug product but, specifically, a compound with at least micromolar affinity for a target [10]. The time devoted to the structure-based drug design process, as outlined in this review, may represent only a fraction of the total time toward developing a marketable drug product. Many years of research may be necessary to convert a drug lead into a drug that will be both effective and tolerated by the human body. Additional years of research and development will bring the drug through clinical trials to finally reach the market.

This review is intended to provide an overview of the process of structure-based drug design from the selection of a target to the generation and evaluation of lead compounds. An in-depth discussion or evaluation of the computational methods involved in drug discovery will not be provided here, since that subject has been covered in reviews elsewhere [11–17].

Overview of the Process

The process of structure-based drug design is an iterative one (see Figure 1) and often proceeds through multiple cycles before an optimized lead goes into phase I clinical trials. The first cycle includes the cloning, purification and structure determination of the target protein or nucleic acid by one of three principal methods: X-ray crystallography, NMR, or homology modeling. Using computer algorithms, compounds or fragments of compounds from a database are positioned into a selected region of the structure. These compounds are scored and ranked based on their steric and electrostatic interactions with the target site, and the best compounds are tested with biochemical assays. In the second cycle, structure determination of the target in complex with a promising lead from the first cycle, one with at least micromolar inhibition in vitro, reveals sites on the compound that can be optimized to increase potency. Additional cycles include synthesis of the optimized lead, structure determination of the new target:lead complex, and further optimization of the lead compound. After several cycles of the drug design process, the optimized compounds usually show marked improvement in binding and, often, specificity for the target.

Choice of a Drug Target

The choice of a drug target is primarily made on a biological and biochemical basis. The ideal target macromolecule for structure-based drug design is one that is closely linked to human disease and binds a small molecule in order to carry out a function. The target molecule

*Correspondence: amy.c.anderson@dartmouth.edu

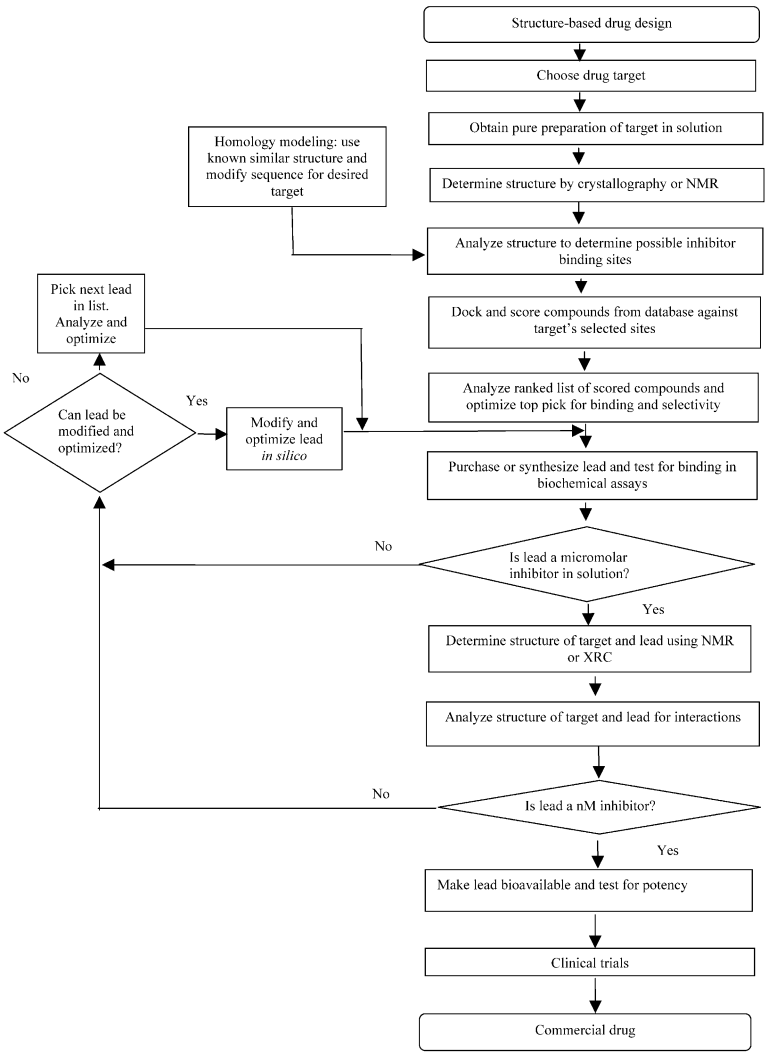

Figure 1. The Iterative Process of Structure-Based Drug Design

usually has a well-defined binding pocket. Other designed small molecules can compete, at a required level of potency, with the natural small molecule in order to modulate the function of the target. Many good drug targets are proteins; however, drug design against RNA targets with well-defined secondary structure, like the bacterial ribosome and portions of the HIV genome, has also been effective. Recent reviews highlight some of the RNA structure-based projects underway [18, 19]. In diseases caused by the malfunction of human proteins, small molecule drugs against G protein coupled receptors (GPCRs) represent at least 25% of the currently marketed drugs [20]. Small molecules that modulate the function of ion channels, proteases, kinases, and nuclear hormone receptors make up another 22% of the market.

The goal in developing drugs against the targets listed above is often to modulate the function of the human protein; the goal in developing drugs against pathogenic organisms is total inhibition, leading to the death of the pathogen. Antimicrobial drug targets should be essential, have a unique function in the pathogen, be present only in the pathogen, and be able to be inhibited by a small molecule. The target should be essential, in that it is part of a crucial cycle in the cell, and its elimination should lead to the pathogen's death. The target should be unique: no other pathway should be able to supplement the function of the target and overcome the presence of the inhibitor. If the macromolecule satisfies all outlined criteria to be a drug target but functions in healthy human cells as well as in a pathogen, specificity can often be engineered into the inhibitor by exploiting structural or biochemical differences between the pathogenic and human forms. Finally, the target molecule should be able to be inhibited by binding a small molecule. Enzymes are often excellent drug targets because compounds can be designed to fit within the active site pocket.

Cancer targets can be difficult because the targets are often somatic cell mutants of proteins that regulate essential cellular functions, resulting in the *loss* of a function. Of course, it is difficult for a small molecule to potentiate the recovery of a function. However, as pointed out in a perspective by W. Kaelin [21], a loss of function in one molecule is often correlated with a gain of function in another. The disruption of oncogenic com-

plexes is another difficult problem for anticancer drug design. For example, a chromosomal translocation in core binding factor β causes the formation of a novel chimeric protein that sequesters necessary transcription factor subunits [22]. Despite the difficulty of designing a small molecule to disrupt an unwanted protein association, the specific interface between the fusion protein and the transcription factor does provide a target that can be exploited. Finally, malignancy often alters the target from its normal behavior, leading to interest in the design of specificity for the malignant state.

Evaluating a Structure for Structure-Based Drug Design

Once a target has been identified, it is necessary to obtain accurate structural information. There are three primary methods for structure determination that are useful for drug design: X-ray crystallography, NMR, and homology modeling. The evaluation of structures from each method will be discussed.

Crystal structures are the most common source of structural information for drug design, since structures determined to high resolution may be available, and the method is useful for proteins that range in size from a few amino acids to 998 kDa [23]. Another advantage of crystallography is that ordered water molecules are visible in the experimental data and are often useful in drug lead design. A crystal structure should be evaluated for the resolution of the diffracted amplitudes (often simply called resolution); reliability, or R factors; coordinate error; temperature factors; and chemical "correctness." Typically, crystal structures determined with data extending to beyond 2.5 Å are acceptable for drug design purposes since they have a high data to parameter ratio, and the placement of residues in the electron density map is unambiguous. The R factor and R_{free} reported for a model are measures of the correlation between the model and experimental data. The R_{free} value should be below 28% and ideally below 25%, and the R factor should be well below 25% in order to use the structure in drug design. If the only structure available for a particular target does not meet the resolution or R factor criteria, drug design projects can still be considered, but the results should be judged carefully.

Low coordinate error in a crystal structure is crucial since van der Waals interactions modulate with the sixth power of the distance between atoms, and directional bonds, such as hydrogen bonds and electrostatic interactions, have a narrow tolerance for both the angle and distance (approximately 0.2 Å). Coordinate error can be measured in many different ways, but two significant methods are the Luzzati method [24], based on averaging coordinate error as a function of R factors that vary with resolution, and methods in which expected errors are calculated based on the temperature factor, or B factor, of an atom and the atom:reflection ratio [25]. The Luzzati coordinate error is often reported in coordinates deposited with the Protein Data Bank (PDB) and should be in the range of 0.2–0.3 Å. For further accuracy in error determination, the B factor (B) and atom:reflection ratio (atom/refl) can be included, as in the Stroud and Fauman method:

$$\text{Expected error} = 0.642 + 0.00852e^{\left(\frac{B}{7.88}\right)} - 0.687 - 0.00223e^{\left(\frac{B}{6.16}\right)}e^{(-2)(atom/refl)}.$$

Temperature factors of atoms in the region of interest should be no greater than the average temperature factor for the molecule. High temperature factors can reflect disorder due to motion of the residue or ligand or a general indication of error, adding to the inaccuracy of atomic positions. In a study reported by Carson et al. [26], the temperature factor was the most highly correlated determinant of R factor. Finally, the molecule should be refined to be consistent with all rules of stereochemical "correctness" known from small molecule structures; deviations from ideal bond lengths should be no greater than 0.015 Å or 3° for bond angles. Planar atoms should be no more than 0.015 Å out of the plane, and there should be no incorrect chiral centers. Finally, at least 90% of the backbone φ and ψ angles should fall into the most favored regions of the Ramachandran plot. The PDB header record lists these statistics, and they should be evaluated before drug design attempts continue. The results of a structure evaluation program, PROCHECK [27], are also available from the PDB and provide additional detail.

Structures determined by nuclear magnetic resonance, using a concentrated protein or nucleic acid in solution, are also valuable sources for drug design. Since the target is in solution, it is sometimes possible to interpret the dynamics of the target from the data [28]. Ensembles of structures are deposited in the PDB, all of which satisfy the distance restraints from the experimental data and show reasonable stereochemical parameters. There is no analogous reliability factor as in crystallography, but the quality of the structure is often measured by the rms deviations of the coordinates of the members of the ensemble from the average structure (often divided into main chain and side positions) and overall stereochemical soundness, including van der Waals violations, phi/psi conformational angle analysis, side chain torsion angle analysis, bond lengths, bond angles, and planarity. NMR data are often collected by measuring nuclear Overhauser effect (NOE) peaks between resonant nuclei that are a distance of less than 5 Å apart in the tertiary structure. Another important statistic for evaluating NMR-derived structures is the number of unfulfilled NOE restraints, otherwise called violations. NOE violations are crosspeaks between resonant nuclei that appear in the experimental data but are unexplained in the model. A final evaluation statistic is the total number of NOE restraints per residue, or data:parameter ratio.

In a survey of 97 deposited NMR structures in the PDB [29], Doreleijers et al. found that the average structure had 11.3 restraints per residue and 61 NOE violations. The precision of the structures, as defined by the circular variance of the backbone dihedral angles, is clearly correlated with the number of restraints per residue. The number of residues in the most favored regions of the Ramachandran plot is also correlated with the number of restraints per residue and a low number of NOE violations. The programs PROCHECK-NMR [30] and WHAT IF [31], the results of which are available from

the PDB, provide additional structure-based details for evaluating NMR structures. One other note to consider is that the average structure from the ensemble may or may not actually exist; therefore, one of the members of the ensemble or the entire ensemble itself may be a better target choice.

If no experimentally determined structure is available, a homology model can be used for drug design [32–34]. To evaluate a homology model, SWISS-MODEL [35] outputs a confidence factor per residue that reflects the amount of structural information used to create that portion of the model. A higher confidence number reflects a lower number of templates and therefore a decreased accuracy. All other methods for judging protein structures, such as stereochemical soundness (bond lengths, bond angles, planarity, and packing) and residues in the most favored regions of the Ramachandran plot, apply to analyzing a homology model as well as to experimentally derived models.

Using the structural information obtained through the above techniques, the structure is then prepared for drug design programs by first adding hydrogen atoms, usually absent in crystal structures determined with data at a resolution lower than 1.0 Å. The protonation and tautomeric states of residues as well as the state of histidine residues (ϵ, Δ, or both nitrogens protonated) should be assigned. Small molecules, such as ions and water molecules, can be included during the lead generation phase in cases where they play structural roles that are crucial for the conformation of the target, otherwise they are usually removed to allow any potential lead to occupy their positions.

Identification of the Target Site

Structure-based design begins with the identification of a potential ligand binding site on the target molecule. Ideally, the target site is a pocket or protuberance with a variety of potential hydrogen bond donors and acceptors, hydrophobic characteristics, and sizes of molecular surfaces. The ligand binding site can be the active site, as in an enzyme, an assembly site with another macromolecule, or a communication site necessary in the mechanism of the molecule. In addition to the well-accepted protein target sites, RNA secondary structural elements can provide excellent target sites since they are species specific, bind ligands, and can be specific for a disease state [36, 37]. Target sites for protein-protein interactions can be difficult to locate since these surfaces are often flat, large, and hydrophobic, but even these difficulties can be surmounted [38–40]. Cocrystallization studies in which the target macromolecule is crystallized with an initial small molecule inhibitor can be invaluable for the determination of a good target site.

Drug Design Methods

Once the structure and target site are identified, there are several paths to developing a good lead based on the structure of the target. These paths can be broadly classified as computer aided versus experimental. Computer-aided methods will be the main focus of this review. An example of an experimental method, by way of contrast, is high-throughput screening with combinatorial chemistry, in which thousands of compounds are tested for biochemical effects.

The computer-aided methods can be further classified into at least three categories: inspection, virtual screening, and de novo generation. In the first category, inspection, known molecules that bind the site, such as substrates or cofactors in the case of enzymes, or peptides in the case of protein:protein or protein:nucleic acid interactions, are modified to become inhibitors based on maximizing complementary interactions in the target site [1, 3, 41, 42]. In virtual screening, databases of available small molecules are docked into the region of interest in silico and scored based on predicted interactions with the site. Finally, for de novo generation small fragments of molecules, such as benzene rings, carbonyl groups, amino groups, etc., are positioned in the site, scored, and linked in silico. The final compounds, created in silico from the linked fragments, then must be synthesized in the laboratory. There is some overlap between the virtual screening and de novo generation classifications. Some programs, for example, LUDI, which is usually used to dock fragments of compounds, are also capable of docking and scoring entire compounds. The programs are classified in Table 1 according to their primary use.

There are many excellent drug design software methods available capable of either virtual screening or de novo generation. This review will focus on a few of the major points necessary to decide on a particular route for lead generation. Extensive reviews of the software are available [11, 12, 14, 15, 43] and are highly recommended for further reading.

Questions that are pivotal in deciding on a method for lead generation are as follows: (1) are molecules available which can be modified to be inhibitors, (2) is there a means for synthesizing novel molecules, and (3) what is the degree of accuracy required at a particular stage of the design process versus the time needed for the calculation? Factors such as the inclusion of protein or ligand flexibility and the effects of solvent increase the time needed for the calculation but also increase the predictive value. Each of these questions will be discussed with reference to available drug design algorithms.

Modifying an Initial Compound

Substrates and cofactors for many proteins have been modified to become excellent inhibitors [1, 3, 41, 42, 44, 45] (see Figure 2 for an example). Initially, the crystal structure is solved in the presence of a substrate, cofactor, or drug lead. Then, modifications to direct the small molecule toward being a potent inhibitor are designed in silico based on the interactions of the molecule with the target site. The newly designed compounds are then scored for binding using evaluative scoring algorithms available in virtual screening methods.

Docking Available Small Molecules versus De Novo Generation

The main advantage to docking compounds from databases such as the Available Chemicals Database (ACD) into the target site is that hit compounds can be purchased and tested using biochemical assays. Alternatively, instead of testing the entire database, a database can be refined to select molecules with a specific motif.

Table 1. Algorithms for Docking Small Molecules or Fragments against a Target

	Program	Flexible Protein?	Flexible Ligand?	Description	Reference
Virtual screening	DOCK	no	yes	docks either small molecules or fragments, includes solvent effects	[46–49]
	FlexX	no	yes	incremental construction	[51]
	FlexE	yes	yes	incremental construction; samples ensembles of receptor structures	[52]
	SLIDE	yes	yes	anchor fragments placed, remainder of ligand added; backbone flexibility	[50]
	Flo98	no	yes	can rapidly dock a large number of ligand molecules, graphically view results	[76]
	ADAM	no	yes	fragments aligned based on hydrogen bonding	[77]
	Hammerhead	no	yes	genetic algorithms to link tail fragments to anchor fragments	[78]
	MCSA-PCR	yes	yes	uses simulated annealing to generate conformations of target	[64]
	AUTODOCK	yes	yes	uses averaged interaction energy grid to account for receptor conformations and simulated annealing for ligand conformations	[79]
	MCDOCK	no	yes	Monte Carlo to sample ligand placement	[80]
	ProDOCK	yes	yes	Monte Carlo minimization for flexible ligand, flexible site	[81]
	ICM	yes	yes	Monte Carlo minimization for protein-ligand docking	[82]
	DockVision	no	no	Monte Carlo minimization	[83]
De novo generation of ligands	LUDI	no	yes	docks and scores fragments	[54]
	GRID	no	yes	calculates binding energies for functional groups	[55]
	MCSS	no	yes	exhaustive search of binding site for functional group minima	[56]
	SMoG	no	yes	knowledge-based scoring function; molecules built by joining rigid fragments	[58]
	CONCERTS	no	yes	fills active site with molecular fragments, links fragments	[57]
	Legend	no	yes	grows molecule atom by atom	[84]
	DLD	no	yes	saturates binding site with sp^3 carbons, later linked	[85]
	GrowMol	no	yes	builds ligands from a library of atom types	[86]
	GenStar	no	yes	builds ligands from sp^3 carbons	[87]
	GROW	no	yes	constructs a peptide by residue addition	[88]
	GroupBuild	no	yes	builds ligand from a predefined library of fragments	[89]
	HOOK	no	yes	searches database of molecular skeletons for fit to binding site; hooks two MCSS functional groups to skeleton	[90]
	SPROUT	no	yes	generates skeletons that fit site, substitutes atoms into skeleton to give molecule with correct properties	[91]
	CAVEAT	no	yes	searches database of small molecules to connect fragments	[92]

Programs such as DOCK [46–49], SLIDE [50], FlexX [51], or FlexE [52] and others (see Table 1) dock databases of compounds and score them according to their interactions with the site. Novel scaffolds for inhibitors can be discovered in this way.

De novo lead generation can give rise to novel compounds; however, it does require a team member who can actually synthesize the intended product in the laboratory. Fragments of molecules, usually small functional groups, are docked into the site, scored, and linked together. Ideally, the fragments can more fully explore the binding site than a predefined compound. Means for predicting the synthetic accessibility of the novel compounds are currently being incorporated in a version of LUDI [14, 53]. LUDI [54], GRID [55], MCSS [56], CONCERTS [57], SMoG [58], and others represent examples of de novo lead generation programs (see Table 1).

Time of Calculation versus Predictive Value

In an initial lead generation run, one common goal is to determine the feasibility of the project and the classes of possible leads that may result. Most programs can be run in a "basic" mode which allows this determination. For instance, DOCK [46–49] can position and score all of the compounds in the ACD quite quickly when run with a single rigid target, rigid ligands, and no solvent modeling. However, the predictive value of these pro-

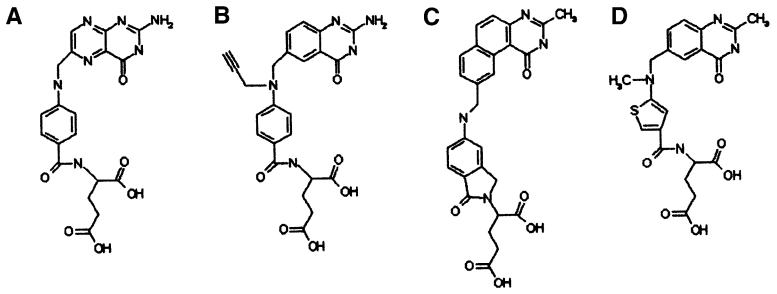

Figure 2. Inhibitors for Thymidylate Synthase Were Designed Based on Modifications of the Cofactor 5,10-Methylene Tetrahydrofolate

Several potent inhibitors are shown: (B) CB3717, (C) OSI 1843U89, and (D) ZD1694 (Tomudex).

grams can be greatly increased when routines that model protein and ligand flexibility as well as solvent contribution are added.

Protein and Ligand Flexibility. There have been many reports which emphasize the crucial effects of including protein and ligand flexibility in the docking and scoring process [15, 43, 59]. Most proteins and most ligands are quite flexible in solution and may experience a full ensemble of possible conformations. As a result, leads generated from a single, rigid structure may have differing results in solution than in silico [60]. In order to account for the landscape of protein and ligand conformations, several drug design algorithms incorporate protein and/or ligand flexibility. However, modeling molecular flexibility, especially for the target macromolecule, drastically increases the compute time required for the structure-based drug design (SBDD) search.

Many programs that allow protein flexibility incorporate information from multiple protein structures. Ensembles of structures can be experimentally determined, such as NMR ensembles (see Figure 3) or multiple crystal structures [61], computationally predicted by molecular dynamics [62], or generated using rotamers of protein side chains [50, 63]. Using a molecular dynamics simulation to generate multiple protein conformations, Carlson et al. have experimentally verified a dynamic pharmacophore model for HIV-1 integrase [62]. Programs which mimic protein flexibility through the use of ensembles include SLIDE [50], FlexE [52], and MCSA-PCR [64].

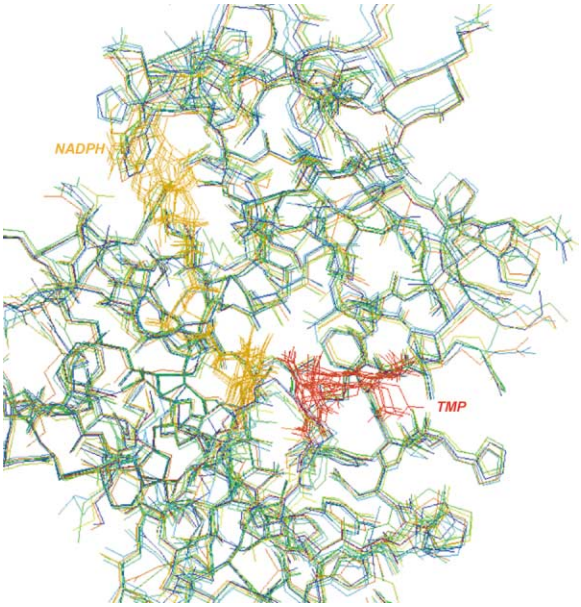

Figure 3. An Ensemble of Six Structures of Dihydrofolate Reductase

Six (out of a total of 24 reported) structures of dihydrofolate reductase bound to trimethoprim (red) and NADPH (orange) (1LUD; [93]) are shown. Each member of the ensemble is separately colored, and hydrogens are omitted for clarity.

Solvent Effects. Solvent plays an important role in ligand binding in several ways. In one capacity, ordered water molecules seen in the structure can be incorporated into the designed ligand, effectively increasing ligand binding by increasing the entropy of the system (releasing the bound water molecule). As an example, inhibitors for HIV protease [65] incorporate an oxygen atom to substitute for a key water molecule coordinated by residues of the flap region of the active site (see Figure 4). In a second capacity, ordered water molecules can be treated as bound ligands, and contacts with them can be maximized [66]. In a third capacity, the effect of the solvent can be incorporated into the scoring scheme for the target:ligand interaction. The steps of increased accuracy in modeling the solvent effect during scoring are as follows: (1) making the assumption that the molecules are in a vacuum, i.e., no solvent modeling; (2) using a fixed dielectric constant in estimating electrostatic contributions; (3) explicit solvation models; and (4) modeling the Born equation. The Born equation calculates the polarization contribution to solvation when a charge is placed within a spherical solvent cavity. In general, increased accuracy comes with increased computational cost.

The correct value for the dielectric constant of the medium is critical in properly evaluating electrostatic effects and estimating binding affinity. In the Northwestern University version of DOCK [49], a solvation correction can be added to the score. Possible approaches to achieve an exact solution to the solvent problem include solving the Poisson-Boltzmann equation, often by using finite differences, or using a free-energy perturbation technique. Three approaches have been used in practice: a modified Born equation [49] to calculate solvation energies, an approximation to the electrostatic desolvation by modeling the first solvation shell at the binding interface [67], and an implicit model which accounts for desolvation by computationally generating possible positions of water molecules in the binding pocket [68].

Drug Lead Evaluation

Once a small molecule has been identified as potentially binding to the target molecule, it must be evaluated

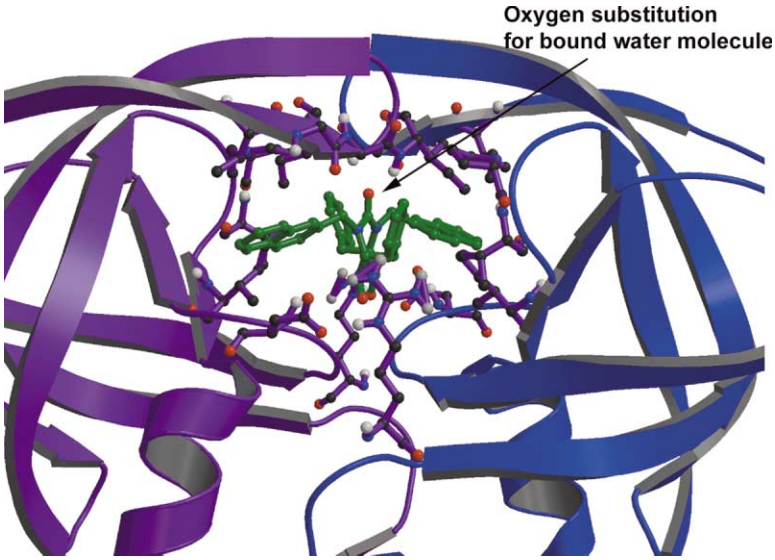

Figure 4. Nonpeptide HIV Protease Inhibitors Based on Cyclic Urea Compounds Incorporate an Oxygen Atom Where a Bound Water Molecule Was Visualized in X-Ray Structures

Nonpeptide HIV protease inhibitors based on cyclic urea compounds incorporate an oxygen atom (noted) where a bound water molecule was visualized in X-ray structures.

before proceeding to further stages. It is important to consider that the ranking assigned by the scoring function is not always indicative of a true binding constant, since the model of the target:ligand interaction is inherently an approximation. Both the solvent effect and the effects of target and ligand flexibility are usually imprecisely described. Usually, several molecules which scored well during the docking run are evaluated in further tests since even the top scoring molecule could fail in vitro assays. Leads are first evaluated visually with computer graphics and can often be optimized at this step for increased affinity. Leads are also evaluated for their likelihood to be orally bioavailable using the "Rule of 5" [69], which states that good leads generally have less than five hydrogen bond donors and less than ten hydrogen bond acceptors, a molecular weight less than 500, and a calculated log of the partition coefficient (clogP) less than 5. Rigidifying the lead can also impart a lower binding constant by decreasing the conformational entropy in the unbound state to approach the presumably very low conformational entropy in the bound state. Veber and colleagues [70] state that the number of rotatable bonds should be less than ten in order to increase the potential for oral bioavailability. Other factors, such as chemical and metabolic stability and the ease of synthesis, can also factor into the decision to proceed with a particular candidate lead. Finally, leads are brought into the wet lab for biochemical evaluation.

Promising leads reenter the structural determination process to find the exact binding mode and to evaluate any further optimization that becomes evident. A few examples of designed leads have shown significant differences between predicted and actual binding modes [71], but in many cases the docked and experimental conformations are within 2 Å rmsd [16].

AmpC β-Lactamase Case Study

There have been many important successes in structure-based drug design.

The discovery of enzyme inhibitors has been successful, since enzymes are often good drug targets and the active site provides an excellent ligand binding site for drug design. Amprenavir (Agenerase) and nelfinavir (Viracept) [72], developed against HIV protease, were designed using mainly structure-based methods and are two of the first drugs to reach the market using SBDD. More recently, zanamivir (Relenza) was developed against neuraminidase [73], Tomudex was developed against thymidylate synthase [44], and imitinab mesylate (Glivec) inhibits Abl tyrosine kinase [74]. With the development of structure-based design against difficult drug targets such as nucleic acids and protein:protein interactions, exciting breakthroughs have recently occurred in the field. Structure-based drug design has revealed specific, micromolar inhibitors against the HIV-1 RNA target TAR [36, 37], the IL-2/IL-2Rα receptor interaction [39], the VEGF/VEGF receptor [40], and Bcl2 [33]. Structure-based design against the enzyme target AmpC β-lactamase illustrates the principles of drug design outlined in this review and will be discussed in further detail in this section.

β-lactamases are bacterial enzymes that cause resistance to β-lactam antibiotics such as the commonly prescribed drugs penicillin and cephalosporin. β-lactamase is a good drug target because it is unique to the pathogen, can be inhibited by a small molecule, and is essential for the pathogen's resistance to β-lactam antibiotics. The β-lactamase enzyme has a serine nucleophile at the active site that cleaves the β-lactam ring of the antibiotic, effectively destroying any pharmaceutical benefit. β-lactamase inhibitors, such as clavulanic acid, are often coadministered with β-lactam antibiotics, but these inhibitors are β-lactams themselves, causing upregulation of the expression of the β-lactamase. Novel β-lactamase inhibitors that do not upregulate expression are needed in order to prevent antibiotic resistance.

The Northwestern University version of DOCK [47, 49] was used to screen the ACD against a consensus structure, a "hot spot" model of AmpC β-lactamase. The consensus structure incorporated experimentally and

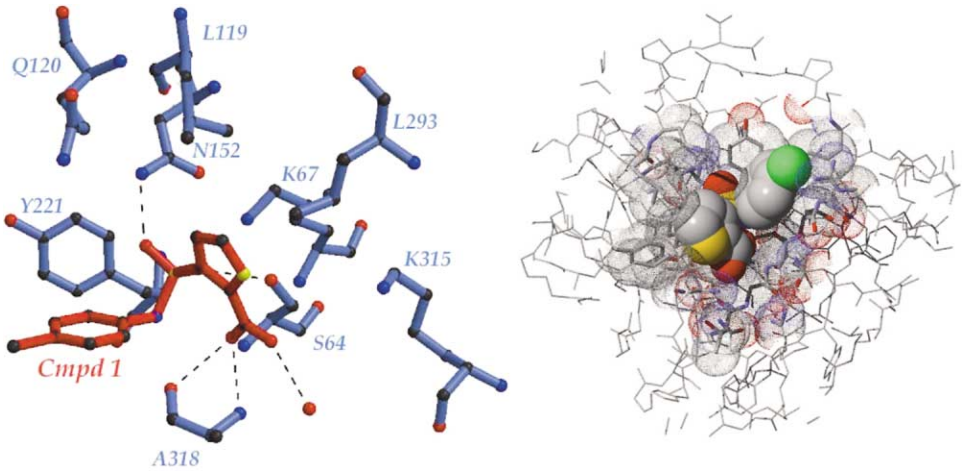

Figure 5. Drug Design against AmpC β-Lactamase

(A) Ball-and-stick representation of compound 1 (red), discovered with a DOCK screen, bound to AmpC β-lactamase.
(B) Compound 1 (space filling) bound to AmpC β-lactamase (residues within 7 Å are shown with van der Waals surfaces).

computationally derived ligand binding data from 13 AmpC β-lactamase structures [75]. The consensus binding sites for AmpC β-lactamase include an amide recognition site, an oxyanion hole, hydroxyl and carboxyl binding sites, and, finally, four ordered water molecules shown to consistently bind either the enzyme or the inhibitors. The top 500 scoring molecules from the DOCK run were examined graphically for complementarity, polar interactions, and agreement with the identified binding sites. Fifty-six compounds were purchased and tested with in vitro assays. Three compounds inhibit with $K_i = 650\ \mu M$ or better. Compound 1 was shown to be selective for AmpC β-lactamase over other serine nucleophile enzymes and was selected for further study.

Powers et al. [66] determined the cocrystal structure of AmpC β-lactamase and compound 1 (Figure 5). The structure was determined to a resolution of 1.94 Å, with R factor 17.3% and R_{free} 20.7%, coordinate error 0.19 Å, average B factor 23 $Å^2$, and average B factor for compound 1, 37 $Å^2$. The structure is stereochemically correct, citing an rmsd from ideality for bond lengths = 0.009 Å and bond angles = 1.5°. The DOCK-predicted conformation of compound 1 closely resembles the crystallographically determined conformation of compound 1. In fact, the rmsd for all inhibitor atoms is 1.87 Å for one molecule in the asymmetric unit of the crystal and 1.75 Å for the second molecule in the asymmetric unit. The predicted interactions were also highly correlated with the crystallographically determined interactions: of nine hydrogen bonds observed in the crystal structure, seven were predicted, and of eight hydrogen bonds predicted, only one was not observed crystallographically.

Compound 1 was tested in microbiology experiments and found to reduce the minimum inhibitory concentration (MIC) of ampicillin by 4-fold in β-lactamase-positive bacteria. Analogs of compound 1 were tested to determine which functional moieties were essential. The carboxylate group, the proton donating ability of the sulfonamide, and the atom order of the sulfonamide were found to be essential. The addition of a piperidine ring to the distal aryl ring increased binding by 2-fold. Finally, compound 1 is relatively "drug-like," according to Lipinski's rules [69], and has sites for future synthetic elaboration.

In summary, AmpC β-lactamase is an excellent drug target with accurate structural information. The Northwestern University version of DOCK was used to screen the ACD to find novel inhibitor scaffolds. The top-scoring compounds were novel and predicted to have complementary interactions with the target site, but were shown to have relatively low binding constants in solution. Further improvement will be needed before the drug lead can proceed into future trials. Structural studies of the selected inhibitor and the enzyme are invaluable in future chemical elaboration.

The results of the AmpC β-lactamase case study also exemplify the sort of reasonable expectations one should have for initial structure-based drug design studies. One, micromolar inhibitors were discovered through the docking procedure and will serve as lead compounds requiring further modification for increased potency. It is very rare that extremely potent inhibitors (nM inhibition or better) are discovered during docking screens. Two, 56 top-scoring compounds were purchased and tested in vitro after the initial docking screen. Due to approximations in the models of protein and ligand interactions in the scoring algorithms, the docked compounds may be ranked in slightly different order than their in vitro assays reveal. In fact, some of the hits from the docking study may not exhibit successful in vitro results at all. Structure-based drug design methods increase the chance that a "hit" will be found in the top-ranked ligands.

Promise for the Future

Structure-based drug design is a powerful method, especially when used as a tool within an armamentarium, for discovering new drug leads against important targets. After a target and a structure of that target are

chosen, new leads can be designed from chemical principles or chosen from a subset of small molecules that scored well when docked in silico against the target. After a preliminary assessment of bioavailability, the candidate leads continue in an iterative process of reentering structural determination and reevaluation for optimization. Focused libraries of synthesized compounds based on the structure-based lead can create a very promising lead which can continue to phase I clinical trials.

As structural genomics, bioinformatics, and computational power continue to explode with new advances, further successes in structure-based drug design are likely to follow. Each year, new targets are being identified, structures of those targets are being determined at an amazing rate, and our capability to capture a quantitative picture of the interactions between macromolecules and ligands is accelerating.

References

1. Roberts, N., Martin, J., Kinchington, D., Broadhurst, A., Craig, J., Duncan, I., Galpin, S., Handa, B., Kay, J., Krohn, A., et al. (1990). Rational design of peptide-based HIV proteinase inhibitors. Science *248*, 358–361.
2. Erickson, J., Neidhart, D., VanDrie, J., Kempf, D., Wang, X., Norbeck, D., Plattner, J., Rittenhouse, J., Turon, M., Wideburg, N., et al. (1990). Design, activity and 2.8 Å crystal structure of a C2 symmetric inhibitor complexed to HIV-1 protease. Science *249*, 527–533.
3. Dorsey, B.D., Levin, R.B., McDaniel, S.L., Vacca, J.P., Guare, J.P., Darke, P.L., Zugay, J.A., Emini, E.A., Schleif, W.A., Quintero, J.C., et al. (1994). L-735,524: the design of a potent and orally available HIV protease inhibitor. J. Med. Chem. *37*, 3443–3451.
4. Mountain, V. (2003). Astex, Structural Genomix, and Syrrx. Chem. Biol. *10*, 95–98.
5. Zheng, D., Huang, Y., Moseley, H., Xiao, R., Aramini, J., Swapna, G., and Montelione, G. (2003). Automated protein fold determination using a minimal NMR constraint strategy. Protein Sci. *12*, 1232–1246.
6. Oezguen, N., Adamian, L., Xu, Y., Rajarathnam, K., and Braun, W. (2002). Automated assignment and 3D structure calculations using combinations of 2D homonuclear and 3D heteronuclear NOESY spectra. J. Biomol. NMR *22*, 249–263.
7. Bailey-Kellogg, C., Widge, A., Kelley, J., Berardi, M., Bushweller, J., and Donald, B. (2000). The NOESY jigsaw: automated protein secondary structure and main-chain assignment from sparse, unassigned NMR data. J. Comput. Biol. *7*, 537–558.
8. Pervushin, K., Riek, R., Wider, G., and Wutrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anistropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc. Natl. Acad. Sci. USA *94*, 12366–12371.
9. Antel, J. (1999). Integration of combinatorial chemistry and structure-based drug design. Curr. Opin. Drug Discov. Dev. *2*, 224–233.
10. Verlinde, C., and Hol, W. (1994). Structure-based drug design: progress, results and challenges. Structure *2*, 577–587.
11. Taylor, R., Jewsbury, P., and Essex, J. (2002). A review of protein-small molecule docking methods. J. Comput. Aided Mol. Des. *16*, 151–166.
12. Joseph-McCarthy, D. (1999). Computational approaches to structure-based ligand design. Pharmacol. Ther. *84*, 179–191.
13. Carlson, H., and McCammon, J. (1999). Method for including the dynamic fluctuations of a protein in computer-aided drug design. J. Phys. Chem. A *103*, 10213–10219.
14. Bohacek, R., and McMartin, C. (1997). Modern computational chemistry and drug discovery: structure generating programs. Curr. Opin. Chem. Biol. *1*, 157–161.
15. Carlson, H., and McCammon, J. (2000). Accommodating protein flexibility in computational drug design. Mol. Pharmacol. *57*, 213–218.
16. Shoichet, B., McGovern, S., Wei, B., and Irwin, J. (2002). Lead discovery using molecular docking. Curr. Opin. Chem. Biol. *6*, 439–446.
17. Klebe, G., and Bohm, H. (1997). Energetic and entropic factors determining binding affinity in protein-ligand complexes. J. Recept. Signal Transduct. Res. *17*, 459–473.
18. Gallego, J., and Varani, G. (2001). Targeting RNA with small-molecule drugs: therapeutic promise and chemical challenges. Acc. Chem. Res. *34*, 836–843.
19. Afshar, M., Prescott, C., and Varani, G. (1999). Structure-based and combinatorial search for new RNA-binding drugs. Curr. Opin. Biotechnol. *10*, 59–63.
20. Hopkins, A., and Groom, C. (2002). The druggable genome. Nat. Rev. Drug Discov. *1*, 727–730.
21. Kaelin, W. (1999). Choosing anticancer drug targets in the post-genomic era. J. Clin. Invest. *104*, 1503–1506.
22. Lukasik, S., Zhang, L., Corpora, T., Tomanicek, S., Li, Y., Kundu, M., Hartman, K., Liu, P., Laue, T., Biltonen, R., et al. (2002). Altered affinity of CBFb-SMMHC for Runx1 explains its role in leukemogenesis. Nat. Struct. Biol. *9*, 674–679.
23. Nissen, P., Hansen, J., Ban, N., Moore, P., and Steitz, T. (2000). The structural basis of ribosome activity in peptide bond synthesis. Science *289*, 920–930.
24. Luzzati, V. (1952). The statistical treatment of errors in crystal structures. Acta Crystallogr. *5*, 802–810.
25. Stroud, R., and Fauman, E. (1995). Significance of structural changes in proteins: Expected errors in refined protein structures. Protein Sci. *4*, 2392–2404.
26. Carson, M., Bugg, C., DeLucas, L., and Narayana, S. (1994). Comparison of homology model to the experimental structure of a novel serine protease. Acta Crystallogr. D Biol. Crystallogr. *50*, 889–899.
27. Laskowski, R., MacArthur, M., Moss, D., and Thornton, J. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.
28. Pellecchia, M., Sern, D., and Wutrich, K. (2002). NMR in drug discovery. Nat. Rev. Drug Discov. *1*, 211–219.
29. Doreleijers, J., Rullman, J., and Kaptein, R. (1998). Quality assessment of NMR structures: a statistical survey. J. Mol. Biol. *281*, 149–164.
30. Laskowski, R., Rullman, J., MacArthur, M., Kaptein, R., and Thornton, J. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR *8*, 477–486.
31. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. *8*, 52–56.
32. Enyedy, I., Lee, S., Kuo, A., Dickson, R., Lin, C., and Wang, S. (2001). Structure-based approach for the discovery of bisbenzamidines as novel inhibitors of matriptase. J. Med. Chem. *44*, 1349–1355.
33. Enyedy, I., Ling, Y., Nacro, K., Tomita, Y., Wu, X., Cao, Y., Guo, R., Li, B., Zhu, X., Huang, Y., et al. (2001). Discovery of small-molecule inhibitors of Bcl-2 through structure-based computer screening. J. Med. Chem. *44*, 4313–4324.
34. Schapira, M., Raaka, B., Samuels, H., and Abagyan, R. (2001). *In silico* discovery of novel retinoic acid receptor agonist structures. BMC Struct. Biol. *1*, 1.
35. Peitsch, M., and Guex, N. (1997). Large-scale comparative protein modeling. In Proteome Research: New Frontiers in Functional Genomics, M. Wilkins, K. Williams, R. Appel, and D. Hochstrasser, eds. (New York: Springer), pp. 177–186.
36. Lind, K., Du, Z., Fujinaga, K., Peterlin, B., and James, T. (2002). Structure-based computational database screening, in vitro assay, and NMR assessment of compounds that target TAR RNA. Chem. Biol. *9*, 185–193.
37. Filikov, A., Mohan, V., Vickers, T., Griffey, R., Cook, P., Abagyan, R., and James, T. (2000). Identification of ligands for RNA targets via structure-based virtual screening: HIV-1 TAR. J. Comput. Aided Mol. Des. *14*, 593–610.
38. Gadek, T., and Nicholas, J. (2003). Small molecule antagonists of proteins. Biochem. Pharmacol. *65*, 1–8.
39. Tilley, J., Chen, L., Fry, D., Emerson, S., Powers, G., Biondi, D.,

Varnell, T., Trilles, R., Guthrie, R., Mennona, F., et al. (1997). Identification of a small molecule inhibitor of the IL-2/IL-2R receptor interaction which binds IL-2. J. Am. Chem. Soc. *119*, 7589–7590.

40. Wiesmann, C., Christinger, H., Cochran, A., Cunningham, B., Fairbrother, W., Keenan, C., Meng, G., and DeVos, A. (1998). Crystal structure of the complex between VEGF and a receptor-blocking peptide. Biochemistry *37*, 117765–117772.
41. Varney, M., Marzoni, G., Palmer, C., Deal, J., Webber, S., Welsh, K., Bacquet, R., Bartlett, C., Morse, C., Booth, C., et al. (1992). Crystal structure-based drug design and synthesis of benz[cd]indole-containing inhibitors of thymidylate synthase. J. Med. Chem. *35*, 663–676.
42. Chan, D., Laughton, C., Queener, S., and Stevens, M. (2001). Structural studies on bioactive compounds. 34. Design, synthesis, and biological evaluation of triazenyl-substituted pyrimethamine inhibitors of *Pneumocystis carinii* dihydrofolate reductase. J. Med. Chem. *44*, 2555–2564.
43. Carlson, H. (2002). Protein flexibility and drug design: how to hit a moving target. Curr. Opin. Chem. Biol. *6*, 447–452.
44. Rutenber, E., and Stroud, R. (1996). Binding of the anticancer drug ZD1694 to *E. coli* thymidylate synthase: assessing specificity and affinity. Structure *4*, 1317–1324.
45. Stout, T., and Stroud, R. (1996). The complex of the anti-cancer therapeutic, BW1843U89, with thymidylate synthase at 2.0 Å resolution: implications for a new mode of inhibition. Structure *4*, 67–77.
46. Kuntz, I., Blaney, J., Oatley, S., Langridge, R., and Ferrin, T. (1982). A geometric approach to macromolecular-ligand interactions. J. Mol. Biol. *161*, 269–288.
47. Lorber, D., and Shoichet, B. (1998). Flexible ligand docking using conformational ensembles. Protein Sci. *7*, 938–950.
48. Ewing, T., Makino, S., Skillman, G., and Kuntz, I. (2001). DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. J. Comput. Aided Mol. Des. *15*, 411–428.
49. Shoichet, B., Leach, A., and Kuntz, I. (1999). Ligand solvation in molecular docking. Proteins *34*, 4–16.
50. Schnecke, V., Swanson, C., Getzoff, E., Tainer, J., and Kuhn, L. (1998). Screening a peptidyl database for potential ligands to proteins with side-chain flexibility. Proteins *33*, 74–87.
51. Kramer, B., Metz, G., Rarey, M., and Lengauer, T. (1999). Ligand docking and screening with FlexX. Med. Chem. Res. *9*, 463–478.
52. Claussen, H., Buning, C., Rarey, M., and Lengauer, T. (2001). FlexE. Efficient molecular docking considering protein structure variations. J. Mol. Biol. *308*, 377–395.
53. Boehm, H. (1996). Towards the automatic design of synthetically accessible protein ligands: Peptides, amides and peptidomimetics. J. Comput. Aided Mol. Des. *10*, 265–272.
54. Boehm, H. (1992). The computer program, LUDI: a new method for the de novo design of enzyme inhibitors. J. Comput. Aided Mol. Des. *6*, 61–78.
55. Goodford, P. (1985). A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J. Med. Chem. *28*, 849–857.
56. Caflisch, A., Miranker, A., and Karplus, M. (1993). Multiple copy simultaneous search and construction of ligands in binding sites: applications to inhibitors of HIV-1 aspartic proteinase. J. Med. Chem. *36*, 2142–2167.
57. Pearlman, D., and Murcko, M. (1996). CONCERTS: dynamic connection of fragments as an approach to de novo ligand design. J. Med. Chem. *39*, 1651–1663.
58. DeWitte, R., and Shakhnovich, E. (1996). SMoG: a de novo design method based on simple, fast, and accurate free energy estimates. 1. Methodology and supporting evidence. J. Am. Chem. Soc. *118*, 11733–11744.
59. Davis, A., and Teague, S. (1999). Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. Angew. Chem. Int. Ed. Engl. *38*, 736–749.
60. Anderson, A., O'Neil, R., Surti, T., and Stroud, R. (2001). Approaches to solving the rigid receptor problem by identifying a minimal set of flexible residues during ligand docking. Chem. Biol. *8*, 445–457.
61. Knegtel, R., Kuntz, I., and Oshiro, C. (1997). Molecular docking to ensembles of protein structures. J. Mol. Biol. *266*, 424–440.
62. Carlson, H., Masukawa, K., Jorgensen, W., Lins, R., Briggs, J., and McCammon, J. (2000). Developing a dynamic pharmacophore model for HIV-1 integrase. J. Med. Chem. *43*, 2100–2114.
63. Lovell, S., Word, J., Richardson, J., and Richardson, D. (2000). The penultimate rotamer library. Proteins *40*, 389–408.
64. Ota, N., and Agard, D. (2001). Binding mode prediction for a flexible ligand in a flexible pocket using multi-conformation simulated annealing pseudo crystallographic refinement. J. Mol. Biol. *314*, 607–617.
65. Lam, P., Jadhav, P., Eyerman, C., Hodge, C., Ru, Y., Bacheler, L., Meek, J., Otto, M., Rayner, M., and Wong, Y. (1994). Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. Science *263*, 380–384.
66. Powers, R., Morandi, F., and Shoichet, B. (2002). Structure-based discovery of a novel, noncovalent inhibitor of AmpC beta-lactamase. Structure *10*, 1013–1023.
67. Majeux, N., Scarsi, M., and Caflisch, A. (2001). Efficient electrostatic solvation model for protein-fragment docking. Proteins *42*, 256–268.
68. Bohm, H. (1998). Prediction of binding constants of protein ligands: A fast method for the prioritization of hits obtained from de novo design or 3D database search programs. J. Comput. Aided Mol. Des. *12*, 309–323.
69. Lipinski, C., Lombardo, F., Dominy, B., and Feeney, P. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. *23*, 3–25.
70. Veber, D., Johnson, S., Cheng, H., Smith, B., Ward, K., and Kopple, K. (2002). Molecular properties that influence the oral bioavailability of drug candidates. J. Med. Chem. *45*, 2615–2623.
71. Fritz, T., Tondi, D., Finer-Moore, J., Costi, M., and Stroud, R. (2001). Predicting and harnessing protein flexibility in the design of species-specific inhibitors of thymidylate synthase. Chem. Biol. *8*, 981–995.
72. Kaldor, S., Kalish, V., Davies, J., Shetty, B., Fritz, J., Appelt, K., Burgess, J., Campanale, K., Chirgadze, N., Clawson, D., et al. (1997). Viracept (Nelfinavir Mesylate AG1343): A potent, orally bioavailable inhibitor of HIV-1 protease. J. Med. Chem. *40*, 3979–3985.
73. Varghese, J. (1999). Development of neuraminidase inhibitors as anti-influenza virus drugs. Drug Dev. Res. *46*, 176–196.
74. Schindler, T., Bornmann, W., Pellicena, P., Miller, W., Clarkson, B., and Kuriyan, J. (2000). Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. Science *289*, 1938–1942.
75. Powers, R., and Shoichet, B. (2002). Structure-based approach for binding site identification on AmpC β-lactamase. J. Med. Chem. *45*, 3222–3234.
76. McMartin, C., and Bohacek, R. (1997). QXP: Powerful, rapid computer algorithms for structure-based drug design. J. Comput. Aided Mol. Des. *11*, 333–344.
77. Mitzutani, M., Tomioka, N., and Itai, A. (1994). Rational automatic search method for stable docking models of protein and ligand. J. Mol. Biol. *243*, 310–326.
78. Welch, W., Ruppert, J., and Jain, A. (1996). Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. Chem. Biol. *3*, 449–462.
79. Goodsell, D., Morris, G., and Olson, A. (1996). Automated docking of flexible ligands: applications of AutoDock. J. Mol. Recognit. *9*, 1–5.
80. Liu, M., and Wang, S. (1999). MCDOCK: A Monte Carlo simulation approach to the molecular docking problem. J. Comput. Aided Mol. Des. *13*, 435–451.
81. Trosset, J., and Scheraga, H. (1999). Prodock: software package for protein modeling and docking. J. Comput. Chem. *20*, 412–427.
82. Abagyan, R., Totrov, M., and Kuznetsov, D. (1994). ICM—a new method for protein modeling and design—applications to docking and structure prediction from the distorted native conformation. J. Comput. Chem. *15*, 488–506.
83. Hart, T., and Read, R. (1992). Proteins *13*, 206–222.
84. Nishibata, Y., and Itai, A. (1993). Confirmation of usefulness of a structure construction program based on three-dimensional

receptor structure for rational lead generation. J. Med. Chem. *36*, 2921–2928.
85. Miranker, A., and Karplus, M. (1995). An automated method for dynamic ligand design. Proteins *23*, 472–490.
86. Bohacek, R., and McMartin, C. (1994). Multiple highly diverse structues complementary to enzyme binding sites: Results of extensive application of a de novo design method incorporating combinatorial growth. J. Am. Chem. Soc. *116*, 5560–5571.
87. Rotstein, S., and Murcko, M. (1993). GenStar: a method for de novo drug design. J. Comput. Aided Mol. Des. *7*, 23–43.
88. Moon, J., and Howe, W. (1991). Computer design of bioactive molecules: a method for receptor-based de novo ligand design. Proteins *11*, 29–34.
89. Rotstein, S., and Murcko, M. (1993). GroupBuild: a fragment-based method for de novo drug design. J. Med. Chem. *36*, 1700–1710.
90. Eisen, M., Wiley, D., Karplus, M., and Hubbard, R. (1994). HOOK: a program for finding novel molecular architectures that satisfy the chemical and steric requirements of macromolecule binding sites. Proteins *19*, 199–221.
91. Gillet, V., Johnson, P., Mata, P., Sike, S., and Williams, P. (1993). SPROUT: a program for structure generation. J. Comput. Aided Mol. Des. *7*, 127–153.
92. Bartlett, P., Shea, G., Telfer, S., and Waterman, S. (1989). CAVEAT: a program to facilitate the structure-derived design of biologically active molecules. In Molecular Recognition: Chemical and Biological Problems, S. Roberts, ed. (Cambridge: Royal Society of Cambridge), pp. 182–196.
93. Polshakov, V., Smirnov, E., Birdsall, B., Kelly, G., and Feeney, J. (2001). NMR-based solution structure of the complex of *Lactobacillus casei* dihydrofolate reductase with trimethoprim and NADPH. J. Biomol. NMR *24*, 67–70.