JOURNAL OF
MEDICINAL
CHEMISTRY

Volume 37, Number 8 April 15, 1994

Perspective

Application of the Three-Dimensional Structures of Protein Target Molecules in Structure-Based Drug Design#

Jonathan Greer,*,† John W. Erickson,‡ John J. Baldwin,§ and Michael D. Varney⊥

Department of Structural Biology, Pharmaceutical Products Division, Abbott Laboratories, Abbott Park, Illinois 60064, Structural Biochemistry Program, PRI/DynCorp, NCI-Frederick Cancer, Research and Development Center, Frederick, Maryland 21702, Department of Medicinal Chemistry, Merck Research Laboratories, West Point, Pennsylvania 19486, and Department of Medicinal Chemistry, Agouron Pharmaceuticals, San Diego, California 92121

Received January 26, 1994

While the chemist of today faces many exciting and stimulating challenges, perhaps the most demanding, promising, and rewarding one is the rational design of novel therapeutic agents for the treatment of human diseases. For many years, the discovery of new drugs has been achieved by taking a lead structure and iterating cycles of new compound synthesis with biological testing of those compounds to derive a structure–activity relationship related to some measure of therapeutic efficacy (Figure 1, black cycle).[1] Initial lead structures either were the natural ligand or were discovered in a random screening program of compounds or fermentation beers using *in vitro* or even *in vivo* tests.[2] Indeed, projects were frequently chosen on the basis of the results of random screening. Sometimes the leads came from literature compounds. Analogs were selected for synthesis during the iteration cycle on the basis of a combination of inspired trial and error and medicinal chemistry experience and intuition.

In recent years, rational drug design has emerged more widely in the pharmaceutical industry.[3] This approach requires selecting a protein target molecule which plays a critical role in a physiologically relevant biological pathway. The chemist typically begins with the natural ligand as the lead and modifies it to produce a compound with the desired properties. The natural ligand or substrate of this protein is manipulated to produce an enzyme inhibitor or an agonist or antagonist for a receptor, depending upon the identified therapeutic need, capitalizing upon knowledge of what is known about the mechanism of action of the protein–ligand complex.

In order to allow the chemist to more fruitfully design modifications of the lead structure, it is helpful to have a three-dimensional structure for the bioactive conformation of the ligand as it binds to the receptor or enzyme. Experience has taught that this conformation is not the solution structure, nor is it the crystal structure of the ligand. Rather it is the conformation of the ligand when it is bound to its receptor or enzyme active site. Knowledge of the bioactive conformation should better permit the chemist to modify analogs constructively to produce novel structures that are potent and specific. A number of methods have been developed to help in the selection and design of better analogs. Quantitative structure activity relations (QSAR),[4] pharmacophore or receptor mapping,[5,6] and more recent 3D QSAR methods, such as CoMFA,[7] have emerged to aid in the discovery of the bioactive conformation and advance the analog design process (Figure 1, blue cycle).

Beyond knowledge of the bioactive conformation of the ligand, it would be valuable to understand the detailed interactions of the ligand with its receptor protein by examining the three-dimensional structure of the protein target in complex with the ligand. This would allow the

This article is based upon an American Chemical Society satellite television presentation entitled "Macromolecular modeling in the discovery of new drugs", broadcast March 6, 1993.

† Department of Structural Biology, Pharmaceutical Products Division, Abbott Laboratories, Abbott Park, IL 60064.

‡ Structural Biochemistry Program, PRI/DynCorp, NCI-Frederick Cancer, Research and Development Center, Frederick, MD 21702.

§ Department of Medicinal Chemistry, Merck Research Laboratories, West Point, PA 19486.

⊥ Department of Medicinal Chemistry, Agouron Pharmaceuticals, San Diego, CA 92121.

0022-2623/94/1837-1035$04.50/0

chemist to preserve the critical interactions with the protein, while modifying the ligand or substrate analog to interact more precisely with the receptor or enzyme and to occupy subsidiary sites, resulting in better potency and specificity. With the structure of the target protein–ligand complex, one can better understand the structure–activity relationships of existing compounds, suggest new analogs to synthesize in current series, and develop novel concepts and ideas for completely new ligand moieties. This methodology is now known as structure-based drug design.

Obtaining the experimental three-dimensional structure of the protein target frequently requires considerable effort to clone the appropriate DNA, express, purify, and characterize multi-milligram amounts of the protein, followed by crystallization and structure solution.[8] For NMR studies, ^{15}N and ^{13}C isotope labeling is frequently crucial to the structure determination.[9] Not only is the three-dimensional structure of the protein target desired, but especially complexes with the lead ligand compounds that are of interest to the chemist.

Sometimes, when sufficient quantities of the protein are not readily available, the structure of the protein can be derived by homology modeling (Figure 1, green cycle).[10,11] The binding of the ligand to the protein can also be examined in this way.[12] The large protein structure database[13] and the exploding protein sequence databases[14–16] are making this approach ever more applicable. While such models are not accurate in the details of their structure, they can provide a good, rapid, albeit approximate, view of the active site and its interactions with the ligand to the chemist to help generate new ideas and designs early in a drug-design project.

Both the ligand-based and protein model-based design strategies can be iterated with chemical synthesis of the designed compounds and biological testing followed by further rounds of ligand or protein model analysis, synthesis, and biological testing (Figure 1, blue and green cycles). However, this process is driven primarily by the experimental data provided by the biological testing. From the structural perspective, one is always extrapolating from the original structural information, and the biological data do not easily permit correcting errors in the structural models.

The existence of an experimental structure for the protein–ligand complex, besides being more accurate, allows one to go beyond examining the details of the binding site and using this information to design new analogs. It allows one to take the designed, synthesized, and assayed compound, return it to the crystal or NMR tube, and redetermine the structure of this complex with this new compound (Figure 1, red cycle). Thus, one can determine experimentally whether the design concept was structurally correct. If the molecule was potent, was it potent for the correct reasons built into the design? If the compound was weaker than expected, how and why did the design concept fail? Best of all, the new structure can be used as the basis for another round of analysis, design, synthesis, and compound testing. This process can be iterated with further rounds of design, synthesis, testing, and so on to ultimately produce potent and specific compounds. This version of the design cycle has proven to be the most powerful implementation of structurally based drug design and is the subject of this review. Close collaboration between the medicinal chemist, molecular modeler/theoretical chemist, pharmacologist, molecular biologist, biochemist, and structural biologist is essential to the optimal utilization of this design cycle.

Three examples are presented here to illustrate how the above ideas are being successfully reduced to practice in a number of important cases leading to the design and synthesis of novel, more potent, and specific compounds than would otherwise have been achieved. The examples illustrate the value of structural information at different stages of the drug-design process. Structural principles and ideas led to the rapid design of a novel class of very potent and highly specific HIV protease inhibitors for the treatment of AIDS. Several iterations of the structure-based drug-design cycle were utilized to optimize the potency of carbonic anhydrase inhibitors for the treatment of glaucoma. The final example describes the *de novo* design, followed by iterative cycles of optimization, of a dramatically new series of thymidylate synthase inhibitors which can serve as anticancer agents. The structure-based drug-design efforts in all of these three cases have led to compounds which are currently in clinical trials.

Design and Structure of C_2 Symmetry-Based Inhibitors of HIV-1 Protease

Background. One of the most active areas of drug discovery research in the world today concerns efforts to stem the tide of the AIDS pandemic. The world-wide search for safe and effective therapies for AIDS has prompted an intensive research effort on the structure and biology of the human immunodeficiency virus (HIV-1) which is the causative agent of this disease. This research had led to the elucidation of a myriad of specific viral targets for drug discovery and design with the result that strategies now exist for targeting virtually every aspect of the viral life cycle.[17] HIV-1 protease (HIV PR) is a virally-encoded enzyme that cleaves, or processes, viral *gag* and *gag-pol* protein precursors during virus assembly and maturation.[18] In 1988, it was observed that deletion mutagenesis of the HIV PR gene resulted in the production of noninfectious, immature virus particles.[19] This experiment demonstrated that HIV PR performs an essential function in the life cycle of HIV and thus makes this enzyme an important target for the design of specific antiviral agents for AIDS.

Structure-based drug design[19] was used to design potent and specific, C_2 symmetry-based inhibitors of HIV PR and to optimize the pharmacologic properties of these compounds. This work ultimately led to the first structure-based clinical compound for this important antiviral target. Additional details of this work can be found in several recent reviews.[20–23] The crystallography of HIV PR is an extremely fast-moving field; there are well over 150 published and unpublished crystal structures of various inhibitor complexes in existence today. Many of these structures as well as complexes with substrate-based and asymmetric peptidomimetic inhibitors have also been reviewed.[24,25]

Structure–Function Considerations for Inhibitor Design. The initial observations that retroviral proteases contain the amino acid triplet, Asp-Thr(Ser)-Gly, and that their proteolytic activity could be inhibited by pepstatin led to the early proposal that these enzymes were related mechanistically to the aspartic protease family of enzymes.[26,27] The latter are bilobal, single-chain enzymes in which each lobe, or domain, contributes an aspartic acid residue to the active site.[28] The active site itself is formed

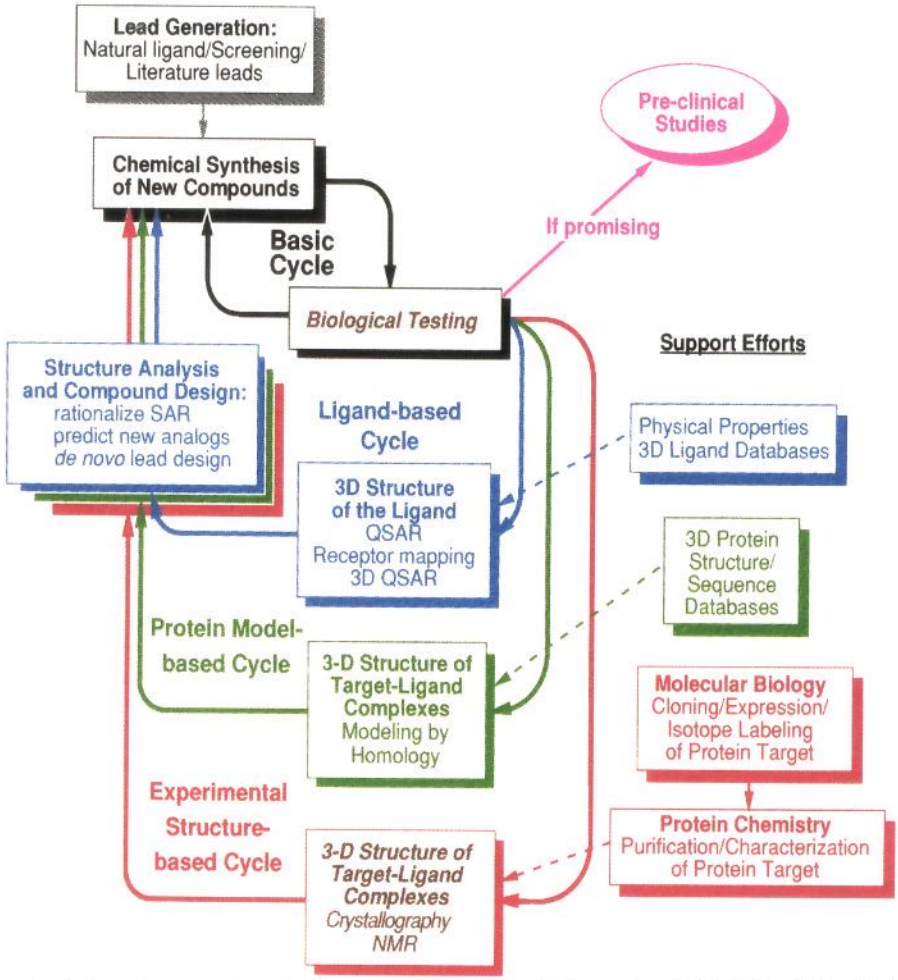


Figure 1. Structure-based drug design cycle. The basic, traditional drug design cycle (labeled "Basic Cycle" in black) begins with a biological assay that tests for the therapeutic use of the compounds. This test is crucial in that it provides the experimental information (labeled in brown) that fuels the cycle. Lead compounds are either the natural ligand or the result of a random screening program or from the literature (highlighted in gray). Analogs are synthesized and tested. On the basis of the testing results, the chemist, using traditional medicinal chemistry, decides what compounds to prepare next, and these compounds are then tested for activity. This basic cycle continues until a satisfactory compound is produced which can be taken for preclinical studies. In structure-based drug design, a number of steps are added progressively as more information becomes available (shown as blue, green, and red cycles). When no information is available about the structure of the receptor, ligand-based strategies, including QSAR, pharmacophore or receptor mapping, and the more recent 3D QSAR methods can be applied (blue cycle).[4-7] On the basis of the results of these methods, structure analysis and compound design are performed which lead to rationalization of existing SAR, new analog prediction, and *de novo* lead design.[87-91] This process is more effective when a model three-dimensional structure can be produced for the receptor protein–ligand complex from known 3D structures and sequences of related proteins using homology modeling methods (green cycle).[10,11] When the experimental crystal or NMR structure of the receptor protein–ligand complex can be determined, then the most powerful structural information can be brought to bear (red cycle). (The figure shows the steps necessary to obtain sufficient protein to perform the experimental structural studies, which frequently includes cloning and expression of the cDNA or gene, and protein purification and characterization.) The three-dimensional structure of the protein–ligand complex is examined in detail to understand the detailed interactions between ligand and protein. On the basis of an analysis of the structure using a variety of increasingly sophisticated computer-assisted drug design strategies and all medicinal chemical knowledge, and with close collaboration of the molecular modeler/theoretical chemist with the medicinal chemist, a new series of analogs is designed. These compounds are synthesized and then tested biologically. The process is iterated until a satisfactory compound is produced for preclinical testing. When an experimental structure determination is available, the iteration process is most effective. The resulting new analogs are recrystallized with the protein or reexamined by NMR to redetermine the experimental structure of the complex of the new analogs with the protein to see whether the design concepts were correct or not. This leads to a new round of compound design, synthesis, and testing. This process is iterated, with cycles of structure determination, analysis, design, synthesis, and biological testing, until a compound with the desired properties is produced. This iterative process is the most powerful since two sets of experimental data fuel the discovery process, the biological testing and the experimental structure determination (both labeled in brown).

at the interface of the N- and C-domains and exhibits approximate 2-fold symmetry at the protein backbone level. Since the sequence length of retroviral proteases is typically about one-third that of aspartic proteases, it was proposed that the former enzymes are composed of two identical subunits, each of which contributes a single aspartic acid to the active site.[29] Crystal structure studies of Rous sarcoma virus (RSV) protease[30] and later of HIV PR by several laboratories[31-34] confirmed that the viral enzymes are homodimers. In the case of the tetragonal crystal form of apo-HIV PR, the dimer exhibits exact crystallographic C_2 symmetry. The active site is formed by the dimer interface and is composed of equivalent contributions of residues from each subunit. The substrate binding cleft is bound on one side by the active-site aspartic acids, Asp25 and Asp125, and on the other by a pair of

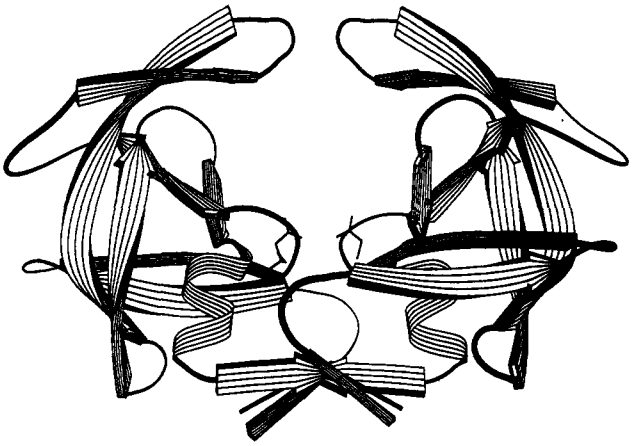


Figure 2. Ribbon drawing of the backbone of HIV-1 protease based on the crystal structure of the native enzyme.[32] The active-site aspartic acid side chains are drawn in stick fashion; the flaps are the two hairpin structures at the top of the molecule. The 2-fold axis of the enzyme is vertical. Adapted from ref 24.

2-fold related, antiparallel β-hairpin structures, or "flaps" (Figure 2). In the crystal structure of RSV protease, the flaps are disordered. In apo-HIV PR, crystal packing forces maintain the flap in a conformation that is presumably unsuitable for substrate binding.[35]

A detailed structural comparison of the retroviral and cellular aspartic proteases revealed that, in contrast to their limited sequence homology, they display considerable structural homology at the backbone level.[36] Fully one-third of the main-chain atoms of RSV PR can be superposed onto the backbone of porcine pepsin to within a 1.5-Å root-mean-square deviation (Figure 3). As expected, most of the structural correspondence is in the active-site region. However, the overall chain topologies of the two families of enzymes are more similar than a simple superposition analysis reveals and are indicative of a distant but definite relationship to a common, ancestral aspartic protease gene.[36]

The close structural and functional relationships between the retroviral and cellular aspartic proteases, together with knowledge of the HIV PR cleavage site sequences, immediately opened the avenue of substrate-based approaches that had been developed for designing inhibitors of renin, an aspartic protease that has long been an important target for the design of antihypertensive agents.[37] Substrate-based inhibitors are essentially peptide substrate analogues in which the scissile peptide bond has been replaced by a noncleavable, transition-state analogue or isostere. This approach has been used to design numerous, highly potent HIV PR inhibitors.[38-41]

Symmetry-Based Inhibitor Design. Despite the enormous collective synthetic effort that has been applied to the design of renin inhibitors, and more recently, HIV PR inhibitors, the usefulness of peptidomimetics as drug candidates has been hampered by their generally poor pharmacologic properties such as oral bioavailability, metabolic stability, and pharmacokinetics.[42] The design of symmetry-based inhibitors for HIV PR represents a significant departure from traditional substrate-based approaches, and one in which knowledge of aspartic protease structure and function could be exploited to conceptualize novel structural classes of inhibitors which, it was hoped, might be more easily developed into potential drug candidates for AIDS.[43,44] Symmetry-based inhibitors had never been designed *a priori* for any enzymatic target, although the concept of symmetry-based inhibitors had been discussed for renin dipeptidase[45] and prostaglandin receptors.[46]

Formulation of the design principles for symmetry-based inhibitors actually began in the absence of knowledge of the structures of either RSV or HIV PR. The hypothesis that HIV PR was a dimeric enzyme, composed of two chemically-identical subunits, led to the postulate that the left and right halves of the active site of this enzyme would be structurally identical, or nearly so. This need not necessarily have been the case; there are numerous examples of symmetric, multisubunit enzymes, but few examples of enzyme active sites that are composed of equivalent, symmetry-related subunits. Similar reasoning led to the construction of a three-dimensional model for HIV PR which embodied exact, C_2 symmetry.[47] If HIV PR incorporates symmetry into its active-site structure, it was reasoned that compounds that would mimic this symmetry might be novel, potent, and specific inhibitors and, furthermore, may be sufficiently non-peptidic in character so as to be pharmacologically superior to the classical peptide-based compounds.

The design strategy had two requirements: first, that the inhibitor possess the same C_2 symmetry as the enzyme, and second, that the symmetry elements of the inhibitor and enzyme approximately superimpose when the inhibitor is bound in the active site.[43,44] Initially a C_2 symmetry-based diaminoalcohol was designed in which a pseudo-C_2 axis passes through the alcohol carbon atom and bisects the O–C–H angle (Figure 4A). Each side of the diamino alcohol resembles a phenylalanine moiety which is a common P_1 substituent for HIV PR substrates. This compound obviously satisfied the first constraint. In order to determine whether the second requirement would be met by this design, a modeling experiment was performed using the crystal structure of a reduced peptide inhibitor complexed with rhizopuspepsin,[48] a fungal aspartic protease, and the crystal structure of RSV PR which had recently been determined and, importantly, made available.[30] The structurally homologous active-site regions of RSV PR and rhizopuspepsin were superimposed in order to "dock" the rhizopuspepsin-bound inhibitor into the active site of RSV PR (Figure 5A). Examination revealed that there were no close contacts. Next, the C-terminal portion of the inhibitor was deleted beyond the reduced CH_2 group, and the N-terminal half was rotated by the enzyme 2-fold axis to produce a pseudo-C_2 symmetric inhibitor (Figure 5B). The deviations from ideal geometry for the computer-generated inhibitor were small enough to suggest that the corresponding diamino alcohol might bind favorably in the orientation as modeled in this experiment. The decision to design symmetric inhibitors with N-terminal properties was based on the experience with renin inhibitors which retain considerable activity after C-terminal truncation.[49]

Symmetry-Based Diamino Alcohols. The prototype compound 1 (Table 1) was synthesized on the basis of the fact that aromatic amino acid side chains are prevalent in the P_1 position of naturally-occurring substrates for HIV PR. This molecule, which closely resembles the central "core" structure of the computer-modeled inhibitor, is pseudosymmetric owing to the secondary OH group on the central carbon atom. Compound 1 was a weak inhibitor of HIV PR (IC_{50} value >200 μM) and did not exhibit significant anti-HIV activity in vitro.[43] Examination of

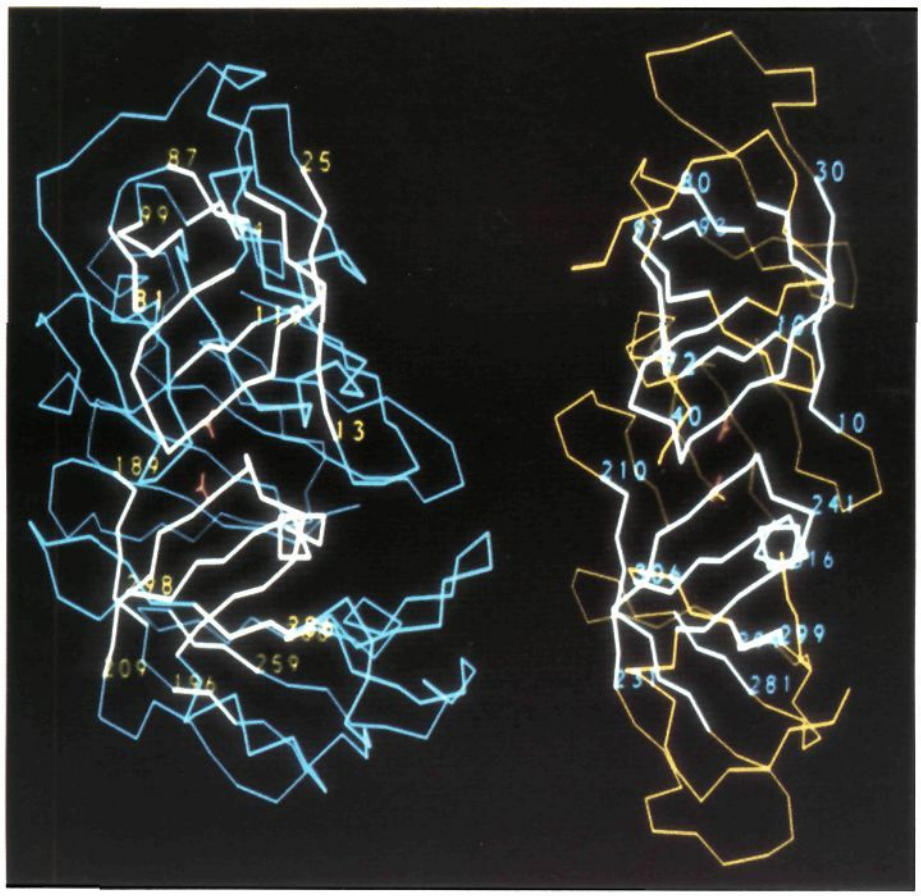


Figure 3. Structural homology of Cα backbones of porcine pepsin (left) and RSV protease (right). Structurally equivalent segments are in white. Adapted from ref 36.

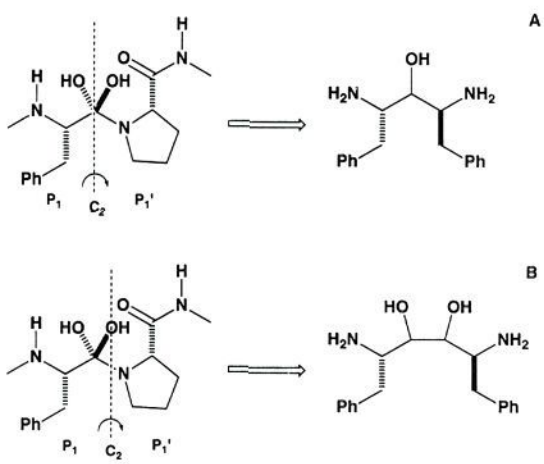


Figure 4. Design of C_2 symmetric inhibitors of HIV-1 protease: (A) placement of the C_2 axis through the carbon atom produces the diamino alcohol; (B) placement of the C_2 axis through the midpoint of the C–N peptide bond produces the diamino diol. Adapted from ref 44.

the substrate binding site region of the modeled structure and reference to the structures of other aspartic protease binding sites suggested that the binding site region of HIV PR should encompass an inhibitor equivalent to at least a hexapeptide in length. These considerations led to the extension of **1** by the symmetric addition of NH_2-blocked amino acids.

The inhibitory potency for a series of diaminoalcohols was measured using a fluorogenic assay[50] and ranged from >10 000 nM for the core structure, **1**, to 3 nM for the bifunctionalized, Cbz-Val compound, **7** (Table 1).[44] Both acetylated and unprotected core compounds, **2** and **1**, respectively, which contain a benzyl moiety in the P_1 position were ineffective inhibitors. Protection with the bulkier Boc group, **3**, resulted in some improvement in potency and suggested a requirement for P_2 substituents. Replacement of Boc by Val, **4**, gave a 5-fold enhancement, and protection of the free amino groups of **4** by acetylation, **5**, resulted in a further 50-fold enhancement. The symmetric addition of acetyl-Val in the P_3 position, **6**, did not yield any improvement over **5**, but substitution of the acetyl group in **5** by Cbz, **7**, resulted in a further 3-fold lowering of the IC_{50} value.

The most potent HIV PR inhibitor of this series, A-74704 (compound **7**), exhibited measurable anti-HIV activity in vitro with an IC_{50} value ≤ 1 μM. Compound **7** also demonstrated good specificity for HIV PR over human renin (>10 000:1), low cellular toxicity (EC_{50}: TC_{50} = 500:1), and was resistant to proteolytic degradation in a renal cortex homogenate at 37 °C ($t_{1/2} \gg 3$ h).[43] Thus, the idea that symmetry-based inhibitors could be specific, potent, and exhibit non-peptide character was partially realized in **7**.

Crystal Structure of Compound 7/HIV PR Complex. To verify that the diamino alcohol inhibitors bound in the predicted symmetric fashion, **7** was cocrystallized with recombinant HIV PR, and the 2.8-Å crystal structure of the complex was solved in the hexagonal space group, $P6_1$.[43] The inhibitor formed a symmetric pattern of hydrogen-bonding interactions with the enzyme (Figure 6) and included a buried water molecule, Wat301, that

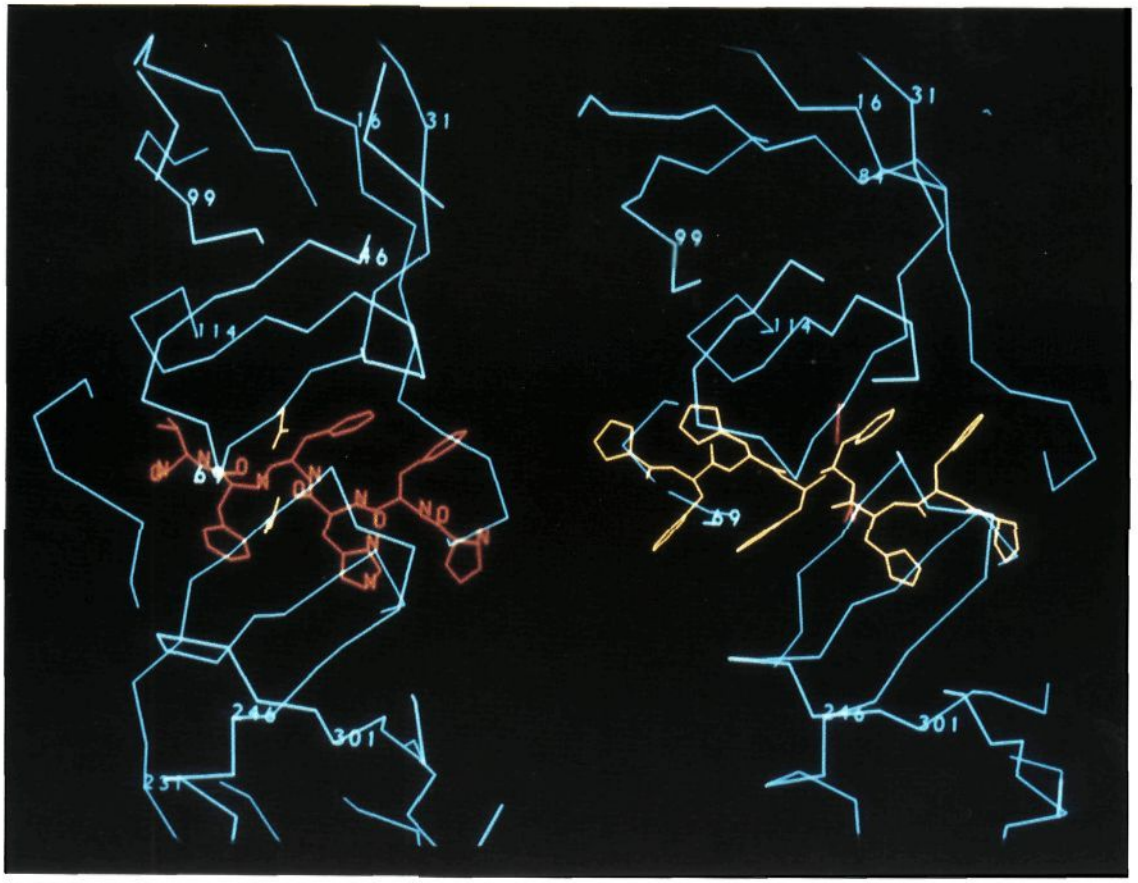
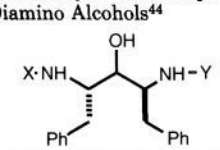


Figure 5. Modeled structures of (A, left) a reduced peptide inhibitor docked into the active site of RSV protease via the superposition of the homologous active site regions of a rhizopuspepsin/inhibitor complex and RSV protease. (B, right) Docked structure of a computer-generated C_2 symmetric inhibitor produced from (A) by C-terminal truncation of the reduced peptide inhibitor beyond the reduced carbon atom and subsequent operation on the remaining N-terminal segment using the dyad of RSV protease. View is down the enzyme dyad. Reprinted with permission.[22]

Table 1. Structure–Activity Relationships for C_2 Symmetry-Based Diamino Alcohols[44]



no.	X	Y	IC_{50} (nM)
1	H-	H–	>10,000
2	Ac-	Ac	>10,000
3	Boc-	Boc	3,000
4	H-Val-	H-Val	590
5	Ac-Val-	Ac-Val	12
6	Ac-Val-Val-	Ac-Val-Val	10
7	Cbz-Val-	Cbz-Val	3

makes bridging hydrogen bonds between the inhibitor P_2 and P_2' CO groups and the enzyme Ile50 and Ile150 NH groups on the flaps. Wat301 is located within 0.2 Å of the enzyme C_2 axis and exhibits approximate tetrahedral coordination. This water is unique to HIV PR and was not predicted from the modeling studies with aspartic proteases. The inhibitor and enzyme C_2 axes pass within 0.2 Å of each other and make an angle of approximately 6°. The symmetry of the inhibitor is nearly exact: 20 non-hydrogen atom pairs from both halves of the inhibitor superimpose to within 0.36 Å rms by an approximate dyad (177.9°). The enzyme 2-fold axis (179.9°) superposed all 99 *Cα backbone atoms of the two subunits to within* 0.42 Å rms after a rotation of 179.9°.

The crystal structure analysis confirmed the proposed symmetry-based mode of binding for the diamino alcohols. This was important since these compounds had been designed on the basis of the RSV PR structure in the absence of knowledge about the flap conformation. Moreover, the crystal structure of native HIV PR indicated that the flaps occupied a position that would not permit them to participate in binding, and it was not easy to predict how the flaps would dock with an inhibitor. The latter question was answered by the determination of the $P2_12_12_1$ crystal structure of a complex of HIV PR with the reduced peptide inhibitor, MVT-101, which revealed that the flaps moved in toward the core by up to 7 Å and assumed a markedly more closed conformation when bound to an inhibitor.[51]

Design of Symmetry-Based Diols. Analysis of the crystal structure of the compound 7/HIV PR complex revealed that the short spacing between the P_1 and P_1' NH groups resulted in hydrogen bonds with poor geometry between the inhibitor P_1 and P_1' amides and the carbonyl groups of Gly27 and Gly127. The replacement of an ethyl alcohol by a glycol isostere has been shown to result in potent renin inhibitors.[52] Thus, a second series of symmetry-based HIV PR inhibitors was designed by application of a C_2 axis placed at the midpoint of the scissile bond (Figure 4B). The resulting diol compounds were generally 10–50-fold more potent than the symmetry-based alcohols.[44] *The additional hydroxymethyl group leads to three* distinct diastereoisomers for the diol analogues. The short,

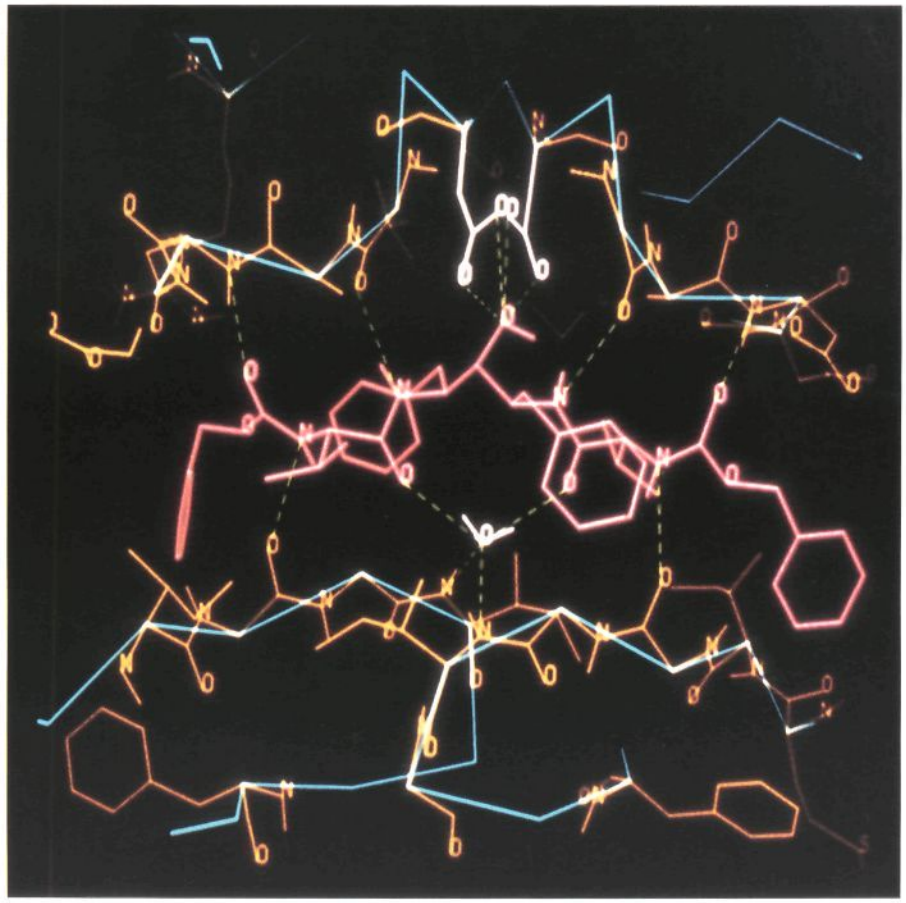
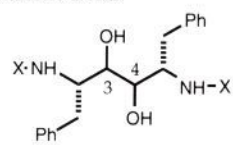


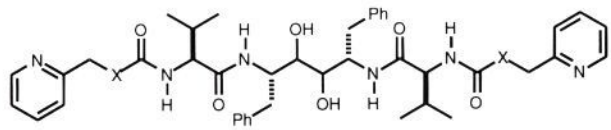
Figure 6. Hydrogen-bonding interactions (dashed lines) between A-74704 (red) and HIV PR. A buried water molecule (white) bridges the inhibitor and the flaps. The active-site carboxylate groups (white) interact with the central hydroxy group on A-74704. Enzyme atoms (yellow) are shown superimposed against the Cα backbone (blue). Reprinted with permission.[22]

Table 2. Structure–Activity Relationships for C_2 Symmetry-Based Diamino Diols[44]



no.	X	conf	IC_{50} (nM)
8	Boc	3*R*,4*R*	40
9	Boc	3*R*,4*S*	12
10	Boc	3*S*,4*S*	280
11	Cbz-Val	3*R*,4*R*	0.22
12	Cbz-Val	3*R*,4*S*	0.22
13	Cbz-Val	3*S*,4*S*	0.38

Table 3. Structure–Activity Relationships and Solubilities for Diamino Diols with Different End-Group Substituents[55]



no.	X	conf	K_i (nM)	EC_{50} (μM)	solubility (pH 7.4, μM)
14	O	3*R*,4*R*	0.16	0.12–0.67	6.5
15	O	3*R*,4*S*	0.09	0.02–0.14	3.1
16	O	3*S*,4*S*	0.19	0.05–0.18	0.27
17	NCH_3	3*R*,4*R*	1.66	0.28–1.5	292
18	NCH_3	3*R*,4*S*	0.15	0.07–0.20	256
19	NCH_3	3*S*,4*S*	0.18	0.06–0.17	4.7

Boc-protected diols, **8–10**, were moderately potent (Table 2). Replacement of Boc with Cbz-Val led to a more potent series, **11–13**, in which the stereochemistry of the two hydroxy groups exhibited surprisingly little effect on inhibitor potency, in sharp contrast to the case for the shorter, Boc-Phe compounds and for substrate-based, hydroxyethylene-containing peptidomimetic inhibitors. The synthesis of the diamino alcohols and diamino diols has been reported.[44,53] Recent and more extensive structure–activity relationships for these compounds have also been reported.[23,54]

The potential usefulness of **11–13** was limited by their poor solubility. Efforts to improve the solubility of this series were aided by examination of the solvent-accessible surface of the compound **7**/HIV PR complex which indicated that the terminal portions of the inhibitors were exposed (Figure 7). Thus, solubility enhancement efforts were directed primarily toward making modifications at the inhibitor termini. This strategy led to a new series of compounds with markedly improved solubilities without sacrificing either enzyme inhibition or antiviral potency (Table 3). A-77003 (compound **18**) was chosen as a clinical candidate and entered phase I/II clinical trials as an intravenous antiviral agent.[55]

Crystal Structure of an (*R*,*S*)-Diol/HIV PR Complex. Modeling analysis of the diols suggested that they could interact with the active-site aspartic acids in two fundamentally different hydrogen bonding configurations (Figure 8). In a symmetric binding mode, each carboxylate group would interact mainly with one hydroxy group, and both hydroxy groups would play nearly equal roles in

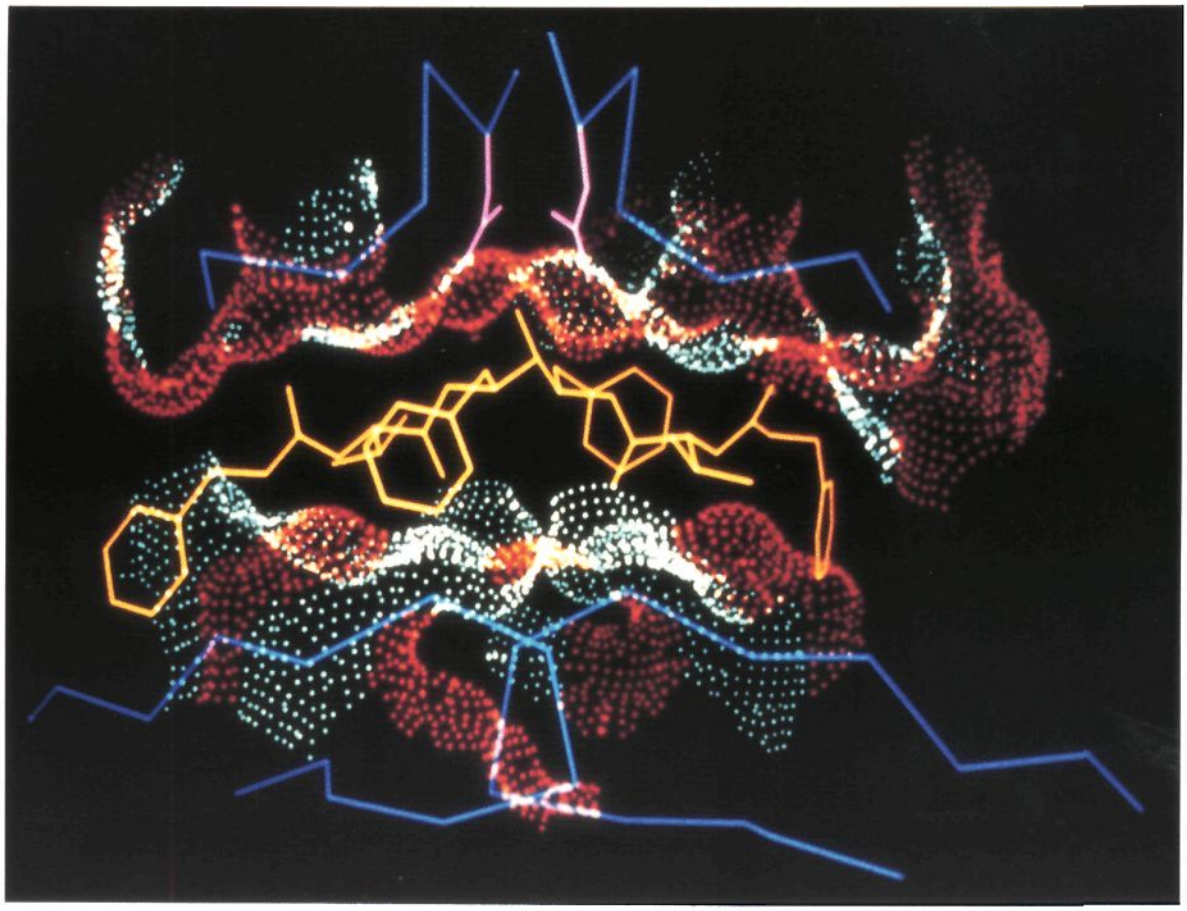


Figure 7. View of a 10–15-Å-thick section of the active site of the A-74704/HIV PR complex. A solvent-accessible surface (dots) of the active-site region was computed after removal of the inhibitor. Color scheme: enzyme backbone (blue); Asp25/125 (pink); A-74704 (yellow); hydrophilic and hydrophobic solvent-accessible surfaces (red and green dots, respectively). Reprinted with permission.[22]

binding. In an asymmetric mode, one of the two hydroxy groups would be located near the enzyme 2-fold axis in a position close to that of the alcohol group of **7** where it could interact with both carboxylate groups. The second hydroxy group would contribute in a minor or perhaps even negative way to the overall interaction energy. Analysis of the crystal structure of the compound **18**/HIV PR complex at 1.8-Å resolution, which was solved in space group $P2_12_12_1$,[56] revealed that this pseudosymmetric, *R,S*-diastereoisomer bound in an asymmetric fashion with respect to the central diol (Figure 9). The (*R*)-OH group on the inhibitor was located between the two active-site aspartic acids, close to the enzyme dyad, where it can form multiple hydrogen bonds to both carboxylate groups. The second (*S*)-OH group is oriented away from the active site and is able to form only a single hydrogen bond. Despite the increased spacing of the P_1–P_1' amides, the geometries of the hydrogen bonds involving Gly27/127 remain poor; one of the distances, 3.8 Å, is too long for an effective hydrogen bond to be formed.

The asymmetric binding observed in the case of **7** could not simply be attributed to the pseudosymmetry of this inhibitor since the pseudosymmetric diamino alcohol, A-74704, bound in a symmetric fashion. Recent crystallographic studies on complexes with two different C_2 symmetric (*R,R*)-diols, compound **11**[25] and SKF-108,361,[57] solved in different space groups, both demonstrated that these compounds bound in an asymmetric fashion. In these complexes, one (*R*)-OH group interacts with both aspartic acids while the second (*R*)-OH group interacts with a single carboxylate group and apparently also with Gly27. Similar results have been observed in a complex with **17**.[56] Thus, even in the case of an inhibitor that can assume an exact C_2 symmetric conformation, asymmetric binding may be observed. These observations suggest that there must be a balance of attractive forces between the central core of the inhibitor and the two aspartic acid residues of the enzyme, on the one hand, and between the peripheral enzyme subsites and the side-chain substituents of the inhibitor, on the other. Preliminary local density functional and semiempirical calculations suggest that the electronic environment surrounding the active-site aspartic acids is asymmetric and this asymmetry may help to promote asymmetric binding modes (I. Topol, S. Burt, and J. Erickson, unpublished). A detailed understanding of the protonation states and charges of the aspartates for different inhibitor complexes is obviously of crucial importance for accurate modeling studies and for quantitative assessments of binding affinities.

Design of Orally Bioavailable Analogs. The diamino diols were designed with the idea that the addition of a second hydroxyl group in the core of the inhibitor would make added hydrogen bonds with the active-site carboxylate groups and that this would lead to an increase in potency over the alcohols. The fact that the (*S*)-OH group in **18** could make only a single additional interaction indicated that this moiety may play a minor or perhaps negative role in binding, since an exposed hydroxyl group in solution may make several hydrogen bonds with water. A-78791, a deshydroxy analogue of **18**, inhibited HIV PR

with greater potency than the parent diol (Figure 10A).[23] To examine the influence of the presence or absence of the (*S*)-OH group on binding of the (*R,S*)-diol, the crystal structure of the complex with A-78791 was solved.[56] Comparison of the two structures indicated that the inhibitors bound in identical conformations, and that the (*S*)-OH in the (*R,S*)-diol had no effect on conformation (Figure 10B). These results strongly suggest that the greater potency of the $CH(OH)CH_2$ analogue over the diols is due to less unfavorable desolvation effects of the former. The greater potency of the diols over the alcohols is still unclear, but may result from a better overall fit of the former in the S′ half of the HIV PR active-site region. Knowledge of the superior potency of the deshydroxy diols led to A-80987, a shorter, orally bioavailable analogue of A-78791 that entered clinical trials.[58]

Summary. The concept of active-site symmetry along with structural and mechanistic considerations of aspartic proteases, has been used to design novel, C_2 symmetry-based inhibitors of HIV PR. Two classes of compounds—a pseudosymmetric diamino alcohol and a diastereomeric set of diamino diols—were initially designed on the basis of the concept of C_2 active-site symmetry. The structure of **7** complexed with HIV PR confirmed the proposed symmetric mode of binding based on initial modeling studies and also proved useful in subsequent efforts to improve the solubility of the more potent diol series. The structure determination of **18** and A-78791 revealed the importance of desolvation effects for inhibitor binding and provided a rationale for the improved potency of the C_2 symmetry-based deshydroxy diol analogues which have been shown to possess superior pharmacological properties, particularly enhanced oral bioavailability.

Symmetry-based inhibitors have subsequently been reported by several groups who have found symmetry to be a useful paradigm for design.[57,59] Symmetry is not a requirement for HIV PR inhibitors as many potent asymmetric substrate-based inhibitors have been designed and several are now in clinical trials.[60] Apparently, the high degree of conformational flexibility of asymmetric peptides and peptidomimetics cancels the structural constraint of having to bind to a symmetric binding cleft. On the other hand, the design of rigid inhibitors, which are often desirable for entropic reasons, should be more tightly constrained by subsite symmetry and will likely require more attention to symmetry to achieve high potency than has been required in the design of peptidomimetic inhibitors. The recent report of cyclized, C_2 symmetric, urea-based inhibitors of HIV PR partly supports this view.[61] While it is not yet possible to extrapolate from structural data to binding potency in a rigorously predictive fashion, it is clear that structure-based approaches to inhibitor design are mature enough to contribute to the conceptualization of medicinal chemistry strategies that can lead to useful clinical candidates for AIDS, cancer, and other diseases.

X-ray Crystallographic-Based Optimization of Carbonic Anhydrase Inhibitors

Background. A long-standing goal of medicinal chemistry has been to move drug discovery from an empirical science toward one based on a structural analysis of ligand–macromolecule interactions. The tools required to achieive this objective of rational design and optimization have become available through advances in molecular biology, macromolecular X-ray crystallography, computer-assisted modeling, and computational analysis.

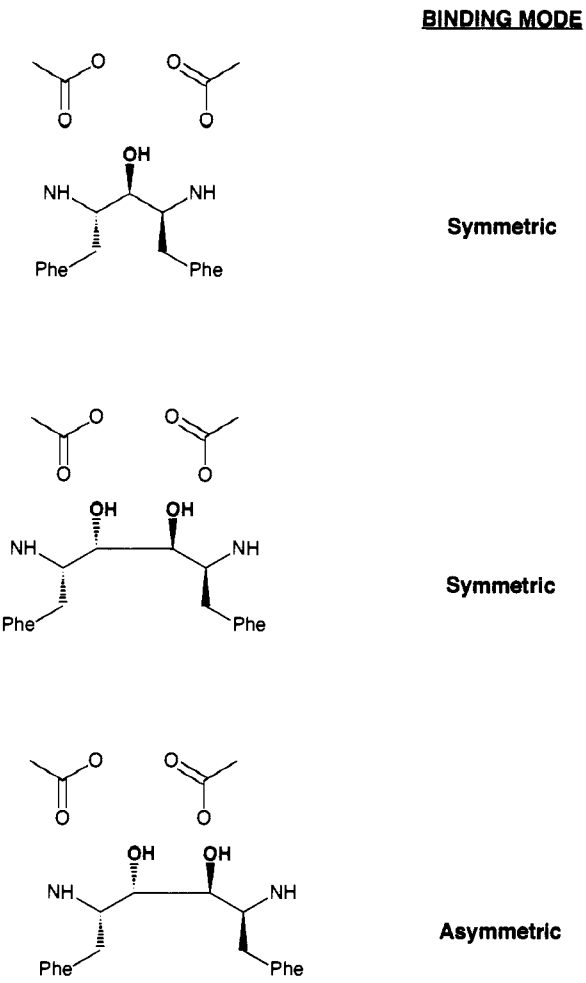


Figure 8. Symmetric vs asymmetric hydrogen-bonding configurations, or binding modes, for symmetry-based diols. The C_2 symmetric (*S,S*)-diol is shown in this example. Reprinted with permission.[56]

The optimization phase of a drug discovery program is especially suitable for an approach based on understanding or rationalizing relative affinities through structural analysis of enzyme–ligand complexes. Such information, coupled with computational chemistry, can define these interactions and suggest further design modification setting the stage for a full iterative cycle.

This iterative approach has been used to facilitate the optimization of carbonic anhydrase inhibitors, the goal being to maximize potency in a series having a proper balance between aqueous solubility and lipophilicity, permitting topical use in the treatment of the ocular disease glaucoma. Such an inhibitor should influence fluid dynamics in the eye through the local blockade of the conversion of carbon dioxide to bicarbonate. This conversion is a critical step in the active secretion of aqueous humor, the fluid that fills the anterior and posterior chambers of the eye. The carbonic anhydrase isozyme found within the secretory cell is carbonic anhydrase II (CA II). Recent evidence suggests that a membrane-bound isozyme, CA IV, may also be involved in the secretory process.[62–64] Both of these isozymes are highly efficient ones being essentially diffusion controlled with a second-order rate constant K_{cat}/K_M of 1.5×10^8 M^{-1} s^{-1} for CA II.

Of the two, CA II is the better characterized in terms of its three-dimensional structure based on X-ray crystal-

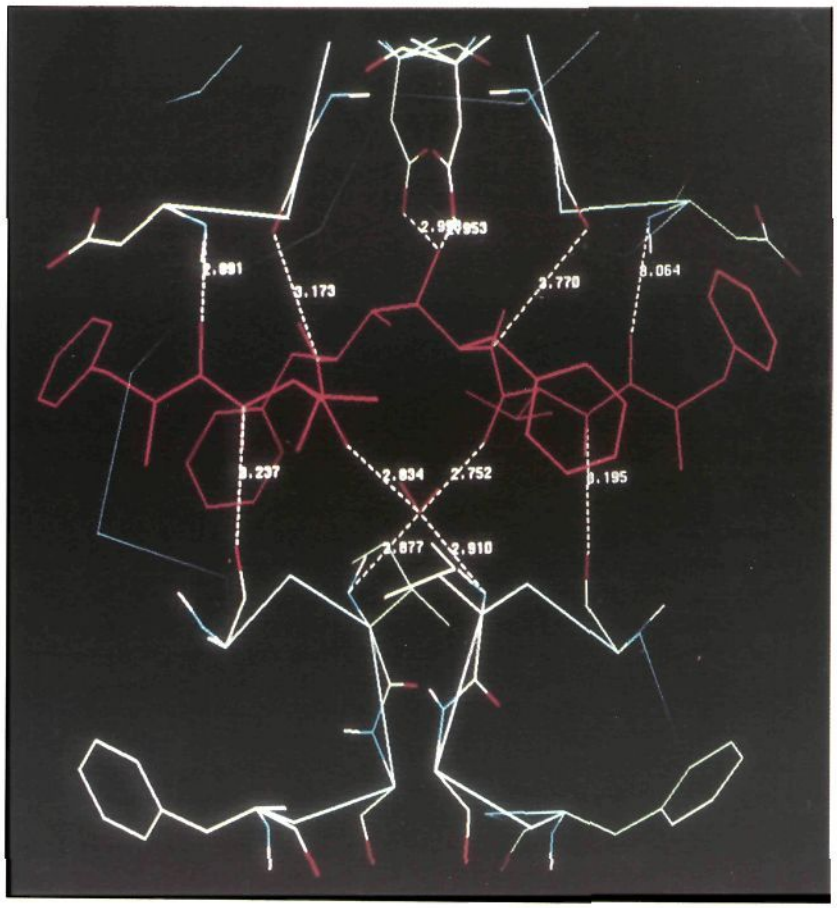
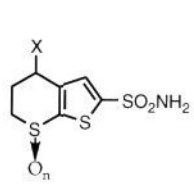


Figure 9. Hydrogen-bonding interactions (dashed lines) between A-77003 and HIV PR. The aspartic acids are on top, the flaps on the bottom, and the inhibitor and buried water molecule (both in red) are sandwiched between them in this view. The (*R*)-OH of A-77003 is shown interacting with both carboxylates; the (*S*)-OH points behind the plane of the figure, away from the aspartates. Numbers refer to hydrogen-bond distances, in angstroms, between heavy atoms (the third decimal place should be ignored). Reprinted with permission.[22]

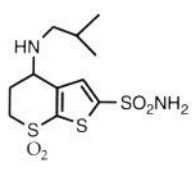
lographic analyses.[65,66] It is this isozyme that served as the model for understanding differences in inhibitory activity based on the structure of enzyme–ligand complexes. Human CA II (HCA II) is a globular enzyme containing 260 amino acid residues with one zinc atom per molecule. The active-site cavity is composed of twisted β sheets being 15 Å across at the surface and descending 16 Å to the zinc atom. The active-site cavity is cone shaped and becomes very narrow at the catalytic hydrophobic pocket where the zinc atom is coordinated in a tetrahedral fashion to three histidine residues: His 94, His 119, and His 96. The active-site cavity is amphiphilic; one wall is dominated by hydrophobic and the other by hydrophilic residues. Both surfaces are important in defining the interaction between ligand and enzyme.[65-70]

Structural Basis for Differences in Inhibitory Potency. The general strategy for the design of a topically effective HCA II inhibitor is illustrated with **20**, a structure which allowed for manipulation of both solubility and lipophilicity. The prototype **21**, MK-927, is water soluble, rapidly penetrates ocular tissue, and lowers intraocular pressure (IOP) in animal models. Resolution provided two enantiomers which differed 100-fold in potency, as determined in a functional enzymatic assay (IC_{50}), and affinity, as measured in a competition assay versus dansyl amide (K_i) (Table 4).

Both the more active *S*- and the less active *R*-enantiomers were cocrystallized with HCA II (Figures 11 and



20

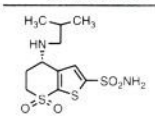
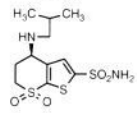


21

12). The X-ray crystallographic results allowed the structure of the bound ligands to be compared with the aim of understanding this 100-fold difference in affinity. With both enantiomers, the sulfonamide group was coordinated to the zinc at the catalytic site through the presumably deprotonated sulfonamide nitrogen while the thiophene ring lay between the hydrophobic and hydrophilic walls of the active-site cavity. Both enantiomers placed the alkyl amino group in the less favored pseudoaxial orientation. *Ab initio* calculations at the 6-31 G* level suggest that the pseudoequatorial conformer would be preferred by about 1 kcal/mol. The potential for hydrogen bonding between both enantiomers and the enzyme involved one oxygen of the SO_2 and the side chain of Glu 92, a polar interaction also was possible between the second oxygen of the SO_2 and the aromatic ring of Phe 131.

The overall geometry of the two inhibitors was similar and analogous to the thiadiazole sulfonamide of acetazolamide. However, two significant differences were found

Table 4

compound	I_{50} (nM) HCAII	K_i (nM) HCA II
	0.54	0.61
S-enantiomer MK-417		
	45	71

between the enantiomers. The first involved the N–S–C–S dihedral angle which was 150° for the *S*-enantiomer and 170° for the *R*. This represents a 20° twist of the thiophene ring in the *R*-isomer relative to the *S*. *Ab initio* molecular orbital calculations at the 3-21 G* (Gaussian 88) level suggest that the preferred dihedral angle is 72°; this implies that the angle formed in the *S*-enantiomer, while not ideal, is preferred by a ΔH of 1 kcal/mol. A second conformational difference between the two enantiomers involved the geometry of the 4-isobutylamino substituent. The side chain is trans in the *S*-enantiomer and gauche in the *R*. Ab initio calculations at the 3–21 G* level suggest that the trans geometry should be preferred by 1 kcal/mol. These two conformational features, differences in the NSCS dihedral angles and the trans vs gauche geometry of the side chain, can account in large measure for the 100-fold difference in affinity and potency observed experimentally for the two enantiomers (Figure 13).[71]

Optimization of Inhibitors. Since accommodation within the active-site cavity requires the isobutylamino group to be oriented in the higher energy pseudoaxial conformation, enhanced affinity could presumably be obtained by decreasing this energic penalty. In one such approach, that of conformation preference, a methyl group was introduced into the 6-position of the thienothiopyran ring system; *ab initio* calculations suggest that this should eliminate the pseudoequatorial preference. To counter the enhanced lipophilicity introduced by the methyl substituent, the 4-isobutylamino group was modified to an ethylamino moiety. All of the four possible optical isomers were prepared and cocrystallized with HCA II, and the structure of each complex was determined by X-ray crystallography. As shown in Table 5, the isomer with the greatest inhibitor potency and affinity had the trans *S,S*-configuration and a K_i value of 0.37 nM. The structural information gained from X-ray crystallography established that the major difference in the conformation of the four optical isomers as bound within the active site was the thiophenesulfonamide NSCS dihedral angle. This torsion angle ranged from 140° for the *S,S*-isomer to 175° for the *R,R* form (Figure 14).

Calculations at the 3–21 G* *ab initio* level suggests this 35° difference amounts to ~1.5 kcal/mol favoring the *S,S*-isomer over its *R,R*-enantiomer. This feature alone accounts for the 20-fold difference found experimentally between the two optical isomers. Likewise, the difference in affinity observed between the two cis isomers can be accounted for by differences in the NSCS dihedral angles.

X-ray crystallographic analysis indicated that the 6-methyl group was immediately above a lipophilic groove, one wall of which is formed by Phe 131. In theory,

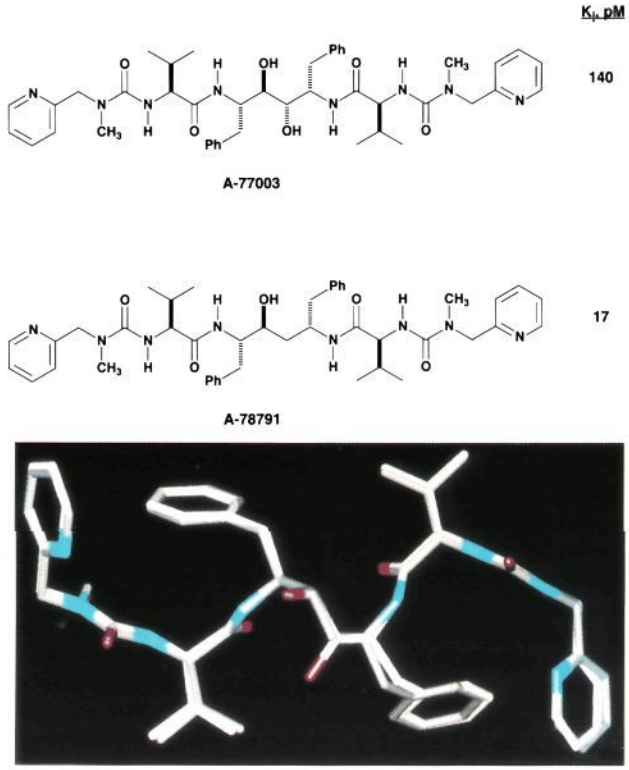
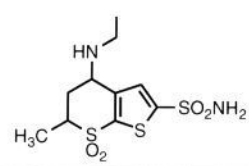


Figure 10. Comparison of A-77003 with its deshydroxy analogue, A-78791: (A, top) chemical structures and inhibition constants for HIV PR; (B, bottom) overlay of inhibitor structures after superposition of the enzyme active sites from the structures of the two inhibitor–enzyme complexes; the two inhibitor structures agree to within 0.15 Å rms for all non-hydrogen atoms.[55] Atoms are color-coded by type.

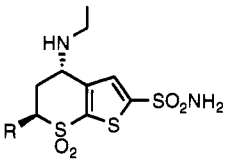
Table 5



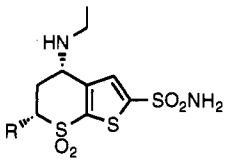
compound	HCA II I_{50} (nM)	HCA II K_i (nM)	dihedral angle (deg) NSCS	ΔH (kcal/mol)
trans *S,S*	0.23	0.37	140	
cis *S,R*	1.1	2.0	153	0.5
trans *R,R*	7.1	5.5	175	1.5
cis *R,S*	29.5	15.3	168	1.5
MK-417 *S*	0.54	0.61	150	0.5

extending the size of this alkyl substituent should increase the nonbonded interaction with the enzyme. As indicated in Table 6, increasing the size of the substituent from methyl to ethyl to propyl results in only a modest improvement in affinity for the trans *S,S*-enantiomers whereas a 10-fold enhancement in K_i was observed with the cis *S,R*-isomers. Modeling studies suggest that failure to improve potency in the trans series is likely due to negative 1,5-synpentane interaction that occurs between the side chain and the lower oxygen of the ring sulfone. Thus, a balance between the increased lipophilic interaction is compensated by a negative steric interaction estimated to be ~1.5 kcal/mol by *ab initio* calculations using a 3–21 G* basis set. With the cis *S,R*-isomers, modeling suggests that the 6-alkyl substituent can engage in nonbonded interactions in the lipophilic groove without unfavorable steric interaction with the sulfone oxygen.

Table 6



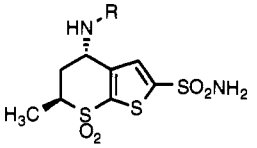
trans S,S



cis S,R

R	K_i (nM) HCA II	R	K_i (nM) HCA II
H_3C	0.37	H_3C	2.0
C_2H_5	0.3	C_2H_5	1.0
C_3H_7	0.14	C_3H_7	0.17

Table 7



R	K_i (nM) HCA II	ΔH (kcal/mol)
H	1.52	
CH_3	1.88	
C_2H_5	0.37	1

In all of the examples studied by X-ray crystallography in this iterative approach, the side chain of His 64 rotated 3.1 Å from its position in the native enzyme. His 64 is a key residue in the enzymatic active site of CA II through its hydrogen-bound waters. This amino acid is believed to be a critical element in the proton shuttle which operates as part of the catalytic mechanism.[72-84] To determine if this movement of the His 64 side chain, induced by the 4-alkylamino group was a positive factor in binding, the size of the substituent was reduced from ethyl to methyl to hydrogen. As shown in Table 7, the 4-amino and 4-methylamino derivatives have 5-fold lower affinity for HCA II than the 4-ethylamine analog. X-ray crystallography established that only the 4-ethylamino example induced movement of His 64, and this movement appeared to be associated with the loss of one ordered water molecule. This suggests that the loss of ordered water from the hydration layer around His 64 and the resulting increase in entropy is responsible for the 5-fold enhancement in affinity found with the larger 6-substituents.

Summary. This iterative approach with multiple X-ray crystallographic analyses established several factors which determine the relative affinity of the thienothiopyran-sulfonamides for HCA II. These include the conformation the inhibitor must adopt in order to be accommodated within the active-site cavity, the polar interactions between the inhibitor and Glu 92, Phe 131, the rotation of the side chain of His 64 with the displacement of an ordered water molecule and lipophilic interactions between the 6-substituent and the surface of the enzyme. Differences in the experimentally determined affinities of these ligands for HCA II can be rationalized by comparing the structures of the enzyme inhibitor complexes in light of these determinants.

Crystal Structure-Based Design of Novel Inhibitors of Thymidylate Synthase

Background. Protein structure-based drug design as it is currently practiced begins with the structure of the target protein solved to atomic resolution using X-ray crystallography. The key to the success of this approach is its iterative nature in that each cycle of design, synthesis, and bioassay is followed by the determination of the high resolution structure of the ligand complexed with the target protein (Figure 1, red cycle). The approach is general and can be applied to any protein or enzyme for which suitable crystals can be grown.

It can be utilized for two distinct purposes. The first, and probably the most powerful, is *de novo* design, essentially the design of completely novel chemical structures from scratch. The structures are designed using the empty active site, created by removing an existing ligand from the protein structure on a graphics machine, and are conceived using the protein structural information as the primary guide. Such a process allows for the design and discovery of structurally unique classes of lead compounds.

The second use of the iterative design process is analogue design or lead optimization. This process begins with the structure of the target protein complexed with a lead compound, sourced from *de novo* design as described above, the chemical or patent literature, or from a random screening of libraries of organic molecules. Analysis of the complex of the lead and target protein results in the design of analogues intended to optimize the chemical or biological properties of the lead. Such properties could include intrinsic enzyme inhibition, plasma protein binding, water solubility, cell permeability, and ease of metabolism. The optimization process is itself iterative and has advantages over traditional structure–activity approaches because the designer knows which portions of the lead interact directly with the protein and which do not. This important information can be used to design modifications to the lead that alter drug properties without compromising inhibition. Taken together, these two design modes result in an integrated lead design and optimization drug-discovery method.

Strategies for *de Novo* Compound Design. Once a particular protein structure has been solved, and before the design process can begin, the active site must be characterized as discussed above. This characterization process is critical to the early phases of this type of discovery program because contained within any protein structure is an enormous amount of information that, if extracted in a systematic fashion, can by exploited for the purposes of design. Both the morphological and electronic properties of the design site are evaluated. This will guide design decisions on the size of substituents, where to place rings, and how to complement discreet or partial charges. In addition, correct orientation of bound water molecules and other hydrogen bonding functionality is established. Finally, a qualitative feeling for the flexibility of active-site side chains and backbone atoms is assessed using temperature factors. Information relating to the flexibility of the protein can greatly influence design decisions particularly with respect to the bulkiness of particular substituents.

During a *de novo* design exercise, filling space is one of the most important properties a novel lead must possess. Once a large portion of the available space is filled by a new skeleton, a hydrogen-bonding functionality can be placed to complement those functional groups presented either by the protein or water molecules. Each bound water molecule is considered individually and may or may not be displaced. Conformational energy calculations are

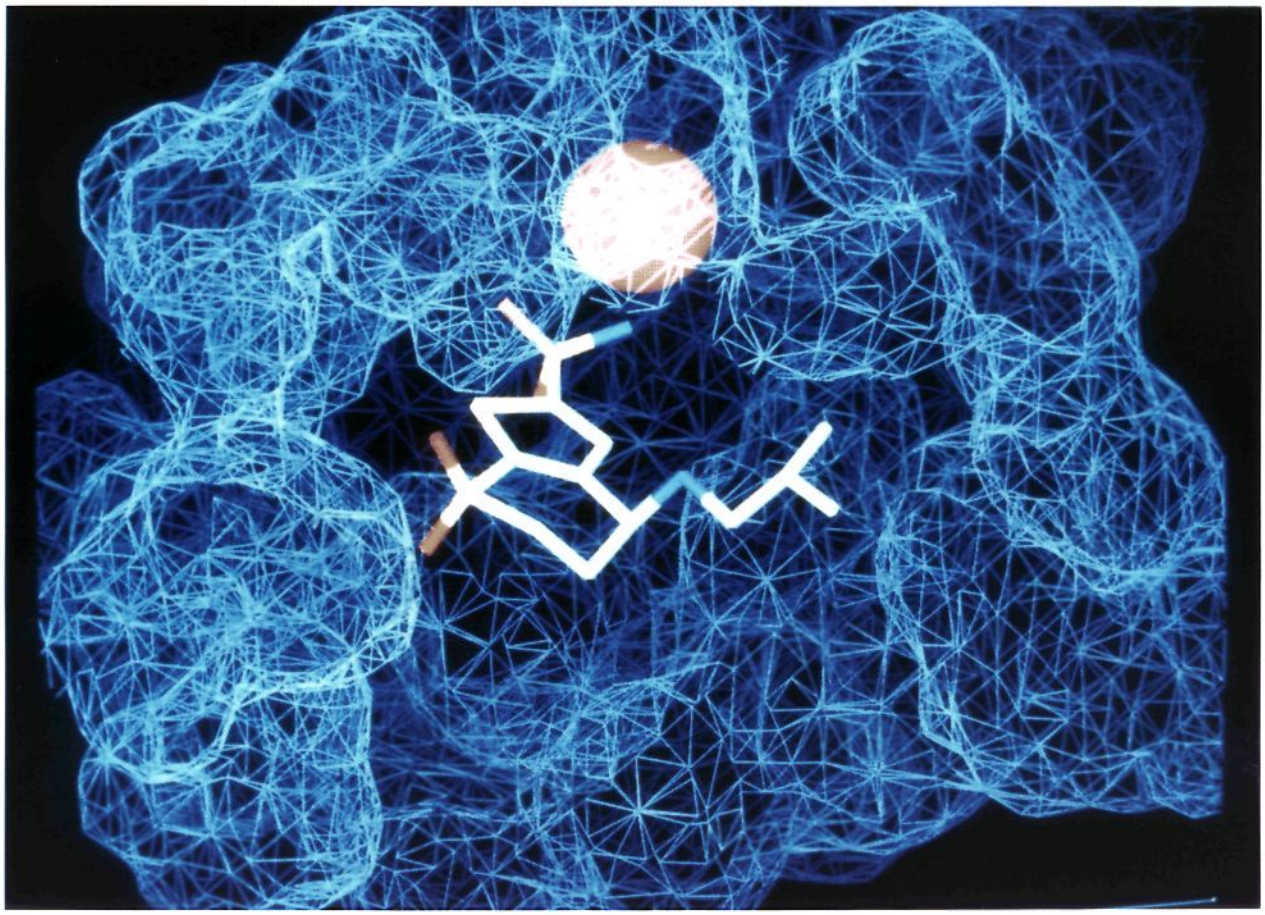


Figure 11. The bound conformation of *S*-2 within the active site of HCA II as determined by X-ray crystallographic studies.

performed on each design idea in order to determine both the internal and desolvation penalties required for the new ligand to attain its binding conformation inside the protein. The internal energy that is required for the small molecule to reach its binding conformation is energy lost in binding. New design ideas are prioritized on the basis of a combination of the above design criteria in addition to ease of synthesis and classical medicinal chemistry principles. The ones with the highest scores are synthesized.

In a structure-based design program, the criteria for defining a lead compound is different from a more traditional approach. In order to be useful, an initial lead compound must have a solubility in the crystallization drop of at least 10 times its inhibition constant. It is more important that a compound be amenable to crystallization with the target protein than that it have a tight binding constant. However, from a practical point of view, compounds with inhibition constants lower than 100 μM are desired. In addition, early compounds in a series should be somewhat flexible and have positions for further elaboration. One does not want to design a lead compound that has all of the positions on a ring substituted and therefore have no room to place substituents during the optimization process.

The design of somewhat flexible leads is highly desired because the positions of protein atoms, even in a highly refined structure, are generally only accurate to 0.2–0.4 Å. Allowing the ligand and the protein to adjust during the binding process can increase the likelihood of optimal overlap. Designated interactions that are placed too rigidly might not be optimal because of even small errors in the positions of atoms in the protein structure. As the design and synthesis cycles progress, rigidity can be built into new ligands once the geometry of desired interactions has been well defined.

***De Novo* Design of a Novel Thymidylate Synthase Inhibitor Series.** The initial area of interest has been in the design of novel inhibitors against the enzyme thymidylate synthase (TS).[85,86] TS is the rate-limiting step is the conversion of deoxyuridylate monophosphate (dUMP) to thymidylate monophosphate. Inhibitors of TS have been shown to be broad-spectrum antiproliferatives and in particular antitumor agents. TS binds two ligands, the substrate dUMP and the cofactor 5,10-methylenetetrahydrofolate. All of the design work which will be described was conducted in the cofactor binding pocket using the structure of the *E. coli* TS complexed with 5-fluoro-2′-deoxyuridylate.

Using a Connolly surface and the methyl probe GRID[87] contours shown in Figure 15 as a guide, a naphthalene ring was positioned into the deep part of the active site such that a maximum overlap occurred between the aromatic ring and the GRID map. Next, the naphthalene ring was used as a scaffold, and hydrogen bonding groups were placed in the 1- and 8-positions to complement an aspartic acid with an NH and a bound water molecule with a carbonyl oxygen. Linkage of the two hydrogen-bonding groups gave a benz[*cd*]indole ring. Empty space off the 6-position of benzindole ring was filled with a dialkylated nitrogen atom. A nitrogen atom was used as the linker atom because it would not create a chiral center

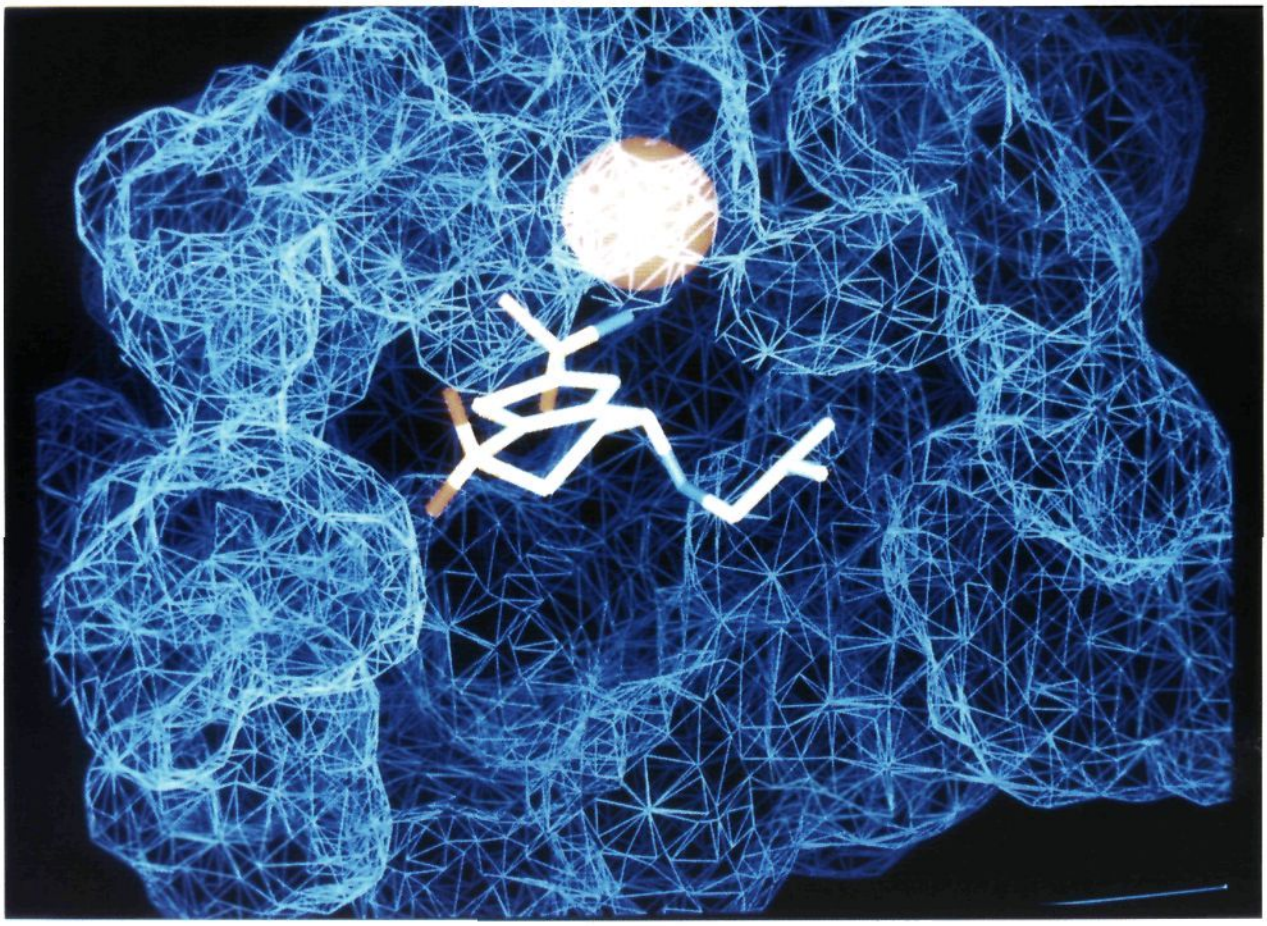


Figure 12. The bound conformation of *R*-2 within the active site of HCA II as determined by X-ray crystallographic studies.

when substituted, and changing its substituents for subsequent designs could be accomplished by simple alkylation chemistry. The fit of this initially designed fragment is shown in Figure 16 superimposed on the methyl probe from GRID. In order to fill the remaining space at the opening of the active site, a benzyl group was used as one of the nitrogen alkyl substituents. To complete the molecule, a solubilizing group (in this case a basic piperazine ring) was placed at the 4-position of the benzene ring. Modeling indicated that the piperazine ring would sit at the interface of protein and bulk solvent. Shown in Table 8 is the initial lead compound **22** along with its inhibition constants against human TS. Solution of the crystal structure of compound **22** complexed with *E. coli* TS was accomplished and is shown in Figure 17. Compound **22** was found to bind in a similar fashion to the original model with one major exception. In the actual fitted structure, the lactam carbonyl oxygen makes an unfavorable interaction with the Ala 263 carbonyl oxygen and as a result moves the Ala 263 oxygen about 1 Å from its previous position. This movement breaks a hydrogen-bond between water 430 and Ala 263 carbonyl oxygen and the water is displaced by the lactam oxygen of the benzindole. It appears that the naphthalene ring wedges itself deeper in the pocket than the model because of more favorable hydrophobic interactions at the top and bottom of the active site. It accomplishes this at the expense of water 430.

Optimization of the New Inhibitor Series. Analysis of this structure resulted in the design of a second generation compound **23** shown in Table 8. Replacement of the lactam carbonyl oxygen with a nitrogen atom was intended to accomplish two goals. The first was that the lactam would be converted to an amidine functionality which should be protonated while bound and therefore make a charged hydrogen bond with Asp 169, and second, the exocyclic NH_2 of the amidine should be an excellent hydrogen-bond donor to the Ala 263 carbonyl oxygen. Since hydrogen-bonding atom distances are inside their respective van der Waals radius, the Ala 263 carbonyl oxygen should be able to move back to its original position and restore the hydrogen bond to water 430. The result would be that the empty space left by the displaced water in the compound **22** complex would be filled. This single atom change resulted in an improvement in binding of roughly a factor of 50 (see Table 8).

The structure of compound **23** complexed with *E. coli* TS is shown in Figure 18. Analysis of this structure revealed the assumptions to be correct. The water 430 had returned and was hydrogen bonded to Ala 263 carbonyl oxygen. The NH_2 of the amidine was hydrogen bonded to both Ala 263 carbonyl oxygen and water 430. In addition, there was a hydrogen bond between the amidine and Asp 169 as predicted.

Using the structure of compound **23** as a guide, a number of derivatives intended to fill remaining space in the deep part of the active site and at the protein solvent interface were designed. Compound **25** has two additional methyl groups on the benzindole ring, both intended to fill the remaining hydrophobic space, and the inhibition constant of this compound is 2 nM against the human enzyme. Compound **24** has replaced the original piperazine ring

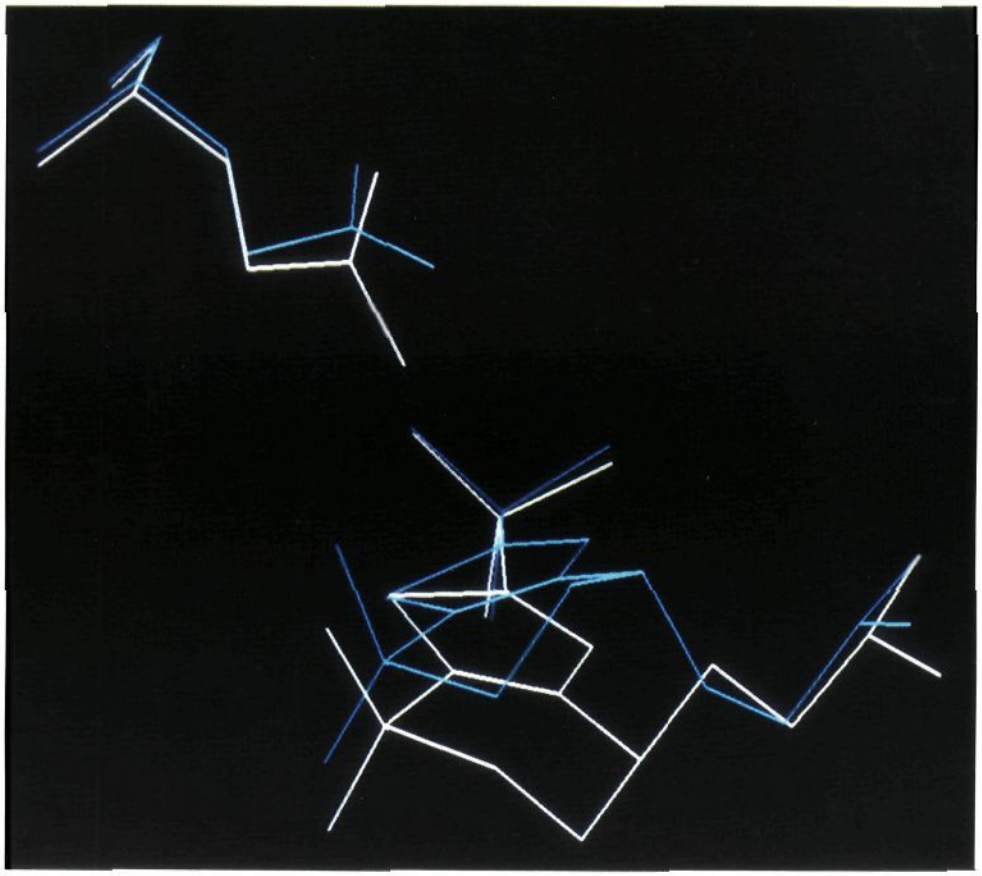
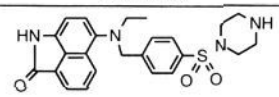
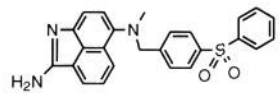
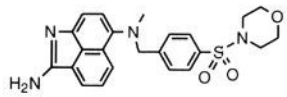
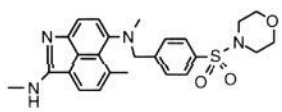


Figure 13. A superposition of the bound conformation of *S*-2 (white) and *R*-2 (blue) illustrating the difference in the N–S–C–S dihedral angle and the gauche versus trans geometry of the amino side chain.

Table 8. Rapid Optimization of a Novel Lead

no.	structure	K_i (μM) human TS	IC_{50} (μM) L1210 cells	thymidine reversal
22		1.6	6.0	1.1
23		0.034	0.38	2.3
24		0.002	0.3	8.5
25		0.002	0.15	12.7

with a morpholine ring and it, too, is a 2 nM inhibitor of human TS. For these two inhibitors, this is an improvement of almost 1000 over the initial lead compound **22**. Table 8 shows selected biological data on the four compounds described. Compound **24** is currently undergoing clinical trials as an antitumor agent.

In summary, an iterative approach to novel inhibitor design has been used that has as its basis structural information gained from X-ray crystallography. *De novo* design of a novel lead and subsequent optimization has resulted in a series of novel and potent benz[*cd*]indole containing TS inhibitors.

Perspectives for the Future

The previous three sections described examples of the success that is currently being achieved using structure-based rational drug design based upon the iterative determination of experimental three-dimensional protein–ligand structures. Since this methodology has been in practice at pharmaceutical companies for at most only 5–10 years, compounds designed using these techniques are only now entering clinical development.

Successes such as the above are fueling the expansion of this still young field and fostering the development of the science and methodology necessary to better design new compounds. An important and rapidly developing part of the structure-based design cycle (Figure 1) is the structure analysis and compound design component. Each of the examples described different strategies used to perform this crucial part of the cycle. There are a broad range of methodologies that have been and are being developed to take advantage of structural information for the design of analogs that are of increasing use and interest to the chemistry community.

It is instructive to consider where the areas of growth and scientific development are in this field. The major purpose of the analysis step in the cycle is the design of the first and subsequent analogs based upon the three-dimensional structure. In the short term, significant efforts are being devoted to the development of better methods for *de novo* lead structure generation, a critical need of the medicinal chemists. *De novo* design is frequently necessary to find a non-peptide lead to mimic or antagonize the actions of a peptide ligand, to identify an alternate series with different pharmacological proper-

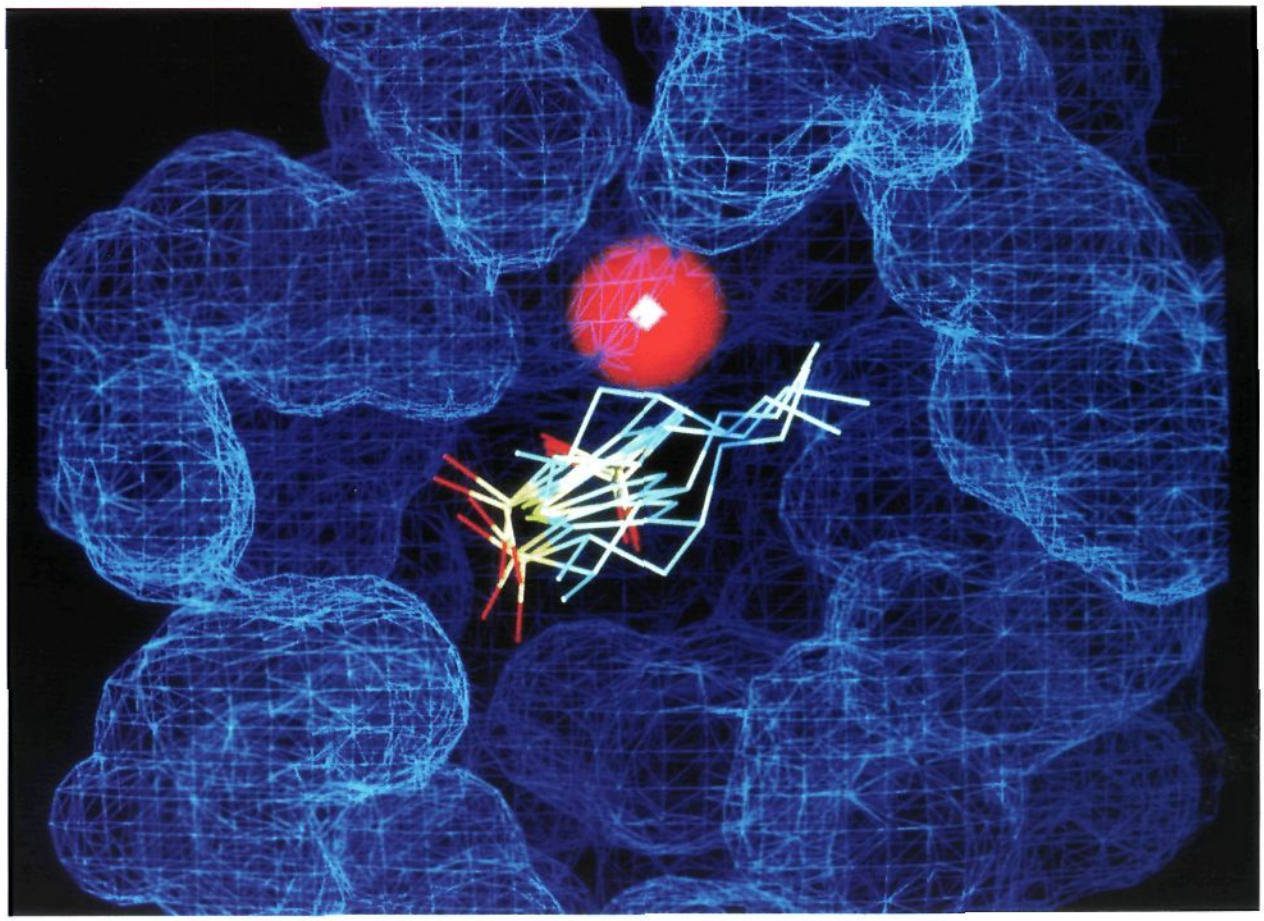


Figure 14. A comparison of the conformation adopted by the four isomers having the 4-(ethylamino)-6-methyl substituents (Table 5) illustrating the differences in the N–S–C–S angle required for accommodation of the inhibitors within the active site cavity of HCA II.

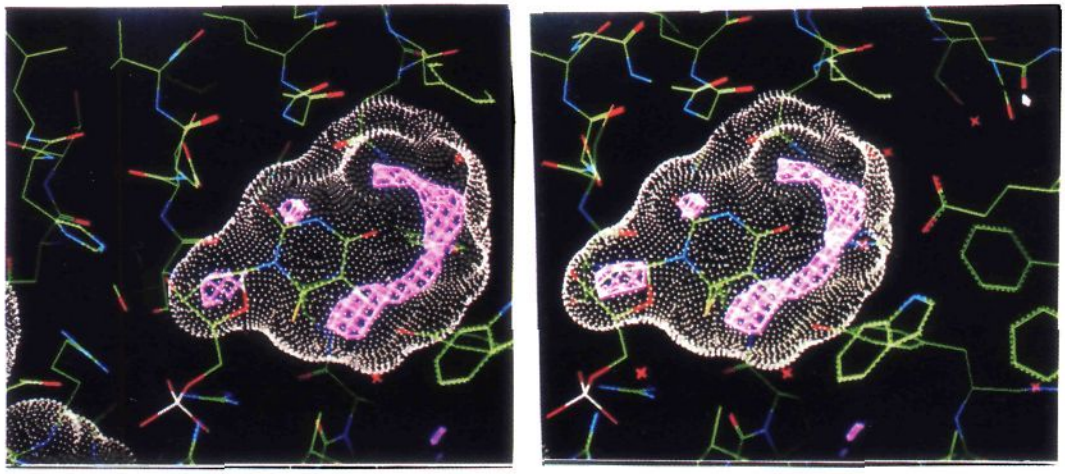


Figure 15. Stereo drawing showing the deep hydrophobic portion of the *E. coli* active site of thymidylate synthase. Shown in purple are the contour maps from GRID, indicating the areas of the active site that are most favorable for a hydrophobic functionality.

ties, or to provide a versatile means for developing a novel series with a consequent strong patent position. Many groups are currently engaged in developing a broad and complementary range of methods to perform *de novo* design based upon the three-dimensional structure of the protein target–ligand complex.[88-91]

Another area of research of considerable importance involves the evaluation of the ligand–protein interactions within the context of a three-dimensional structure of the complex. Better methods are needed to help predict new analogs to be synthesized on the basis of the structural information, providing answers to such questions as: Where are the sites for additional methyl groups that will give a 10-fold increase in potency? Which hydrogen-bonding properties need to be satisfied to improve selectivity? What substituent will properly fill a particular hydrophobic pocket?

Central to the above problem of analog and lead design, indeed to full utilization of the structural information, is the proper evaluation of the free energy of binding of the

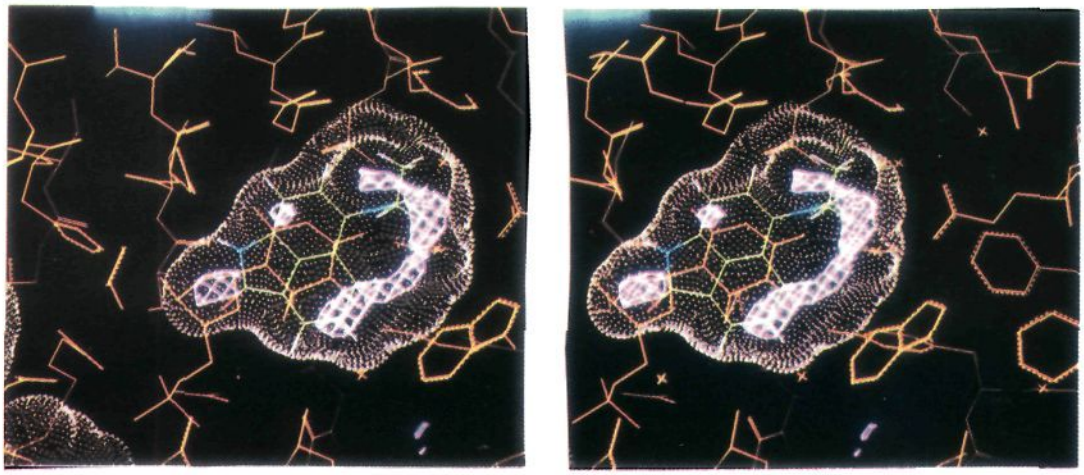


Figure 16. Stereodrawing showing the structure of the initially designed benz[*cd*]indole ring overlaid onto the GRID contours in the active site of TS.

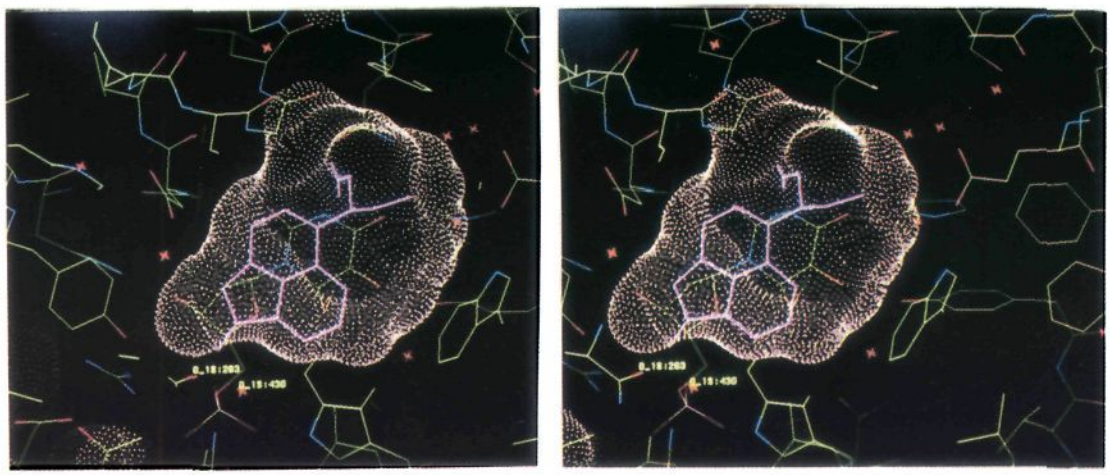


Figure 17. Stereodrawing showing the crystal structure of compound **22** complexed with *E. coli* TS. The position of water 430 has been displayed for illustrative purposes only.

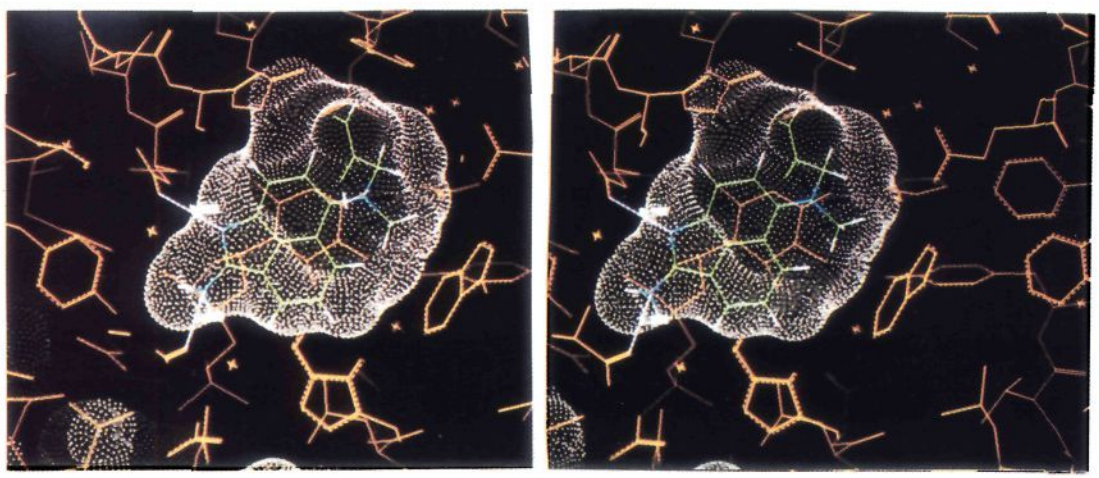


Figure 18. Stereo drawing showing the crystal structure of compound **23** complexed with *E. coli TS*. Important hydrogen bonds are indicated with white dashed lines.

ligand to the protein enzyme or receptor. Since both inhibition and antagonism are generally directly related to the energetics of binding of the ligand, calculation of this quantity is crucial to structure-based drug design. The limitation that currently exists in the ability to correctly calculate such energies is what makes it so essential today to use an iterative scheme requiring redetermination of the experimental structure of the protein–ligand complex, such as was described and utilized in the previous sections. If the conformation and associated energetics of binding could be properly computed, experimental structures would not be so necessary. Much effort is being expended in many laboratories to improve our ability to calculate these free energies to predict binding constants and thereby inhibitory or antagonist potencies.[92-99] When it is possible to perform this calculation

in a reliable manner, structure-based drug design will become an indispensable part of the drug discovery process.

Looking further in the future, the advent of the human genome project is beginning to provide an explosion of sequence information on new potential protein targets for drug discovery. The enormous rate at which these sequences are being and will be developed suggests that experimental structure determinations for these proteins will lag years to decades behind the sequence data. It is crucial therefore to expand on the ability to derive the three-dimensional structure of a protein from its sequence. Unfortunately, the general protein folding problem is likely to remain difficult to solve in the foreseeable future.[100-104] However, homology methods[10,11] do exist for constructing model structures for proteins by extrapolation from experimentally derived structures of related proteins with homologous, or similar, sequences. These methods are likely to become increasingly more important to provide three-dimensional structural information rapidly from sequences. While this can be done at the present time, the structures are in general only approximate, and many of the structural details are not quite correct, limiting its usefulness in drug design, where accurate details are often critical.[12] Significant advances are needed to properly improve and refine these models. Required are both accurate energy evaluation of the protein structures, especially the conformations of external loop segments,[105-112] as well as robust strategies to deal with the multiple minimum problem,[113-116] i.e., that there exist many local minima to the energy function of the protein and it is therefore very difficult and time consuming to find the correct one. Proper refinement procedures, however, would make these models sufficiently accurate for drug-design purposes and greatly expand the possible targets in which structure-based drug design would be an effective aid to the medicinal chemist.

The current successes reported here and elsewhere in the literature, combined with the promise of even more exciting results as the field advances, further indicate that structure-based drug design is a method that is likely to have a significant impact on rational drug design and provide new and better therapeutic agents more effectively and more efficiently in the foreseeable future.

Acknowledgment. J.W.E. acknowledges the collaborative efforts and support of numerous colleagues at Abbott Laboratories and at the NCI-Frederick Cancer Research & Development Center who have made important contributions to this work. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organization imply endorsement by the U.S. Government.

References

(1) Propst, C. L.; Perun, T. J. Introduction to Computer-Aided Drug Design. In *Computer-Aided Drug Design*; Perun T. J., Propst, C. L., Eds.; Marcel Dekker, Inc.: New York, 1989; pp 1–18.
(2) Black, J. Drugs from emasculated hormones: the principle of syntopic antagonism. *Science* **1989**, *245*, 486–493.
(3) Perun, T. J.; Propst, C. L. *Computer-Aided Drug Design*; Marcel Dekker, Inc.: New York, 1989; pp 493.
(4) Martin, Y. C.; Danaher, E. B.; May, C. S.; Weininger, D.; Van Drie, J. H. Strategies in Drug Design Based on 3D-Structures of Ligands. In *QSAR: Quantitative Structure-Activity Relationships in Drug Design*; Alan R. Liss: New York, 1989; pp 177–181.
(5) Martin, Y. C. Overview of Concepts and Methods in Computer-Assisted Rational Drug Design. *Methods Enzymol.* **1991**, *203*, 587–613.
(6) Marshall, G. R.; Barry, C. D.; Bosshard, H. E.; Dammkoehler, R. A.; Dunn, D. A. The conformational parameter in drug design: the active analog approach. In *Computer-Assisted Drug Design, American Chemical Society Symposium*; Olson, E. C., Christoffersen, R. E., Eds.; ACS: Washington, 1979; Vol. 112, pp 205–226.
(7) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *170*, 5959–5967.
(8) Abraham, D. J. X-Ray Crystallography and Drug Design. In *Computer-Aided Drug Design*; Perun, T. J., Propst, C. L., Eds.; Marcel Dekker, Inc.: New York, 1989; pp 93–132.
(9) Fesik, S. W. Approaches to Drug Design Using Nuclear Magnetic Resonance Spectroscopy. In *Computer-Aided Drug Design*; Perun, T. J., Propst, C. L., Eds.; Marcel Dekker, Inc.: New York, 1989; pp 133–184.
(10) Greer, J. Comparative model-building of the mammalian serine proteases. *J. Mol. Biol.* **1981**, *153*, 1027–1042.
(11) Greer, J. Comparative Modeling Methods: Application to the Family of the Mammalian Serine Proteases. *Proteins* **1990**, *7*, 317–334.
(12) Hutchins, C.; Greer, J. Comparative modeling of proteins in the design of novel renin inhibitors. *Crit. Rev. Biochem. Mol. Biol.* **1991**, *26*, 77–127.
(13) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. J.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The protein data bank: a computer-based archival file for macromolecular structures. *Arch. Biochem. Biophys.* **1978**, *185*, 584–591.
(14) Bilofsky, H. S.; Burks, C.; Fickett, J. W.; Goad, W. B.; Lewitter, F. I.; Rindone, W. P.; Swindell, C. D.; Tung, C. S. The GenBank genetic sequence data bank. *Nucleic Acid Res.* **1986**, *14*, 1–4.
(15) George, D. G.; Barker, W. C.; Hunt, L. T. The Protein Identification Resource (PIR). *Nucleic Acids Res.* **1986**, *14*, 11–16.
(16) Hamm, G. H.; Cameron, G. N. The EMBL data library. *Nucleic Acids Res.* **1986**, *14*, 5–10.
(17) Mitsuya, H.; Yarchoan, R.; Broder, S. Molecular targets for AIDS therapy. *Science* **1990**, *249*, 1533–1544.
(18) Debouck, C. The HIV-1 protease as a therapeutic target for AIDS. *AIDS Res. Hum. Retroviruses* **1992**, 8, 153–164.
(19) Kohl, N. E.; Emini, E. A.; Schleif, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A. F.; Scolnick, E. M.; Sigal, I. S. Active human immunodeficiency virus protease is required for viral infectivity. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4686–4690.
(20) Kempf, D. J. Design of symmetry-based, peptidomimetic inhibitors of HIV protease. *Methods Enzymol.*, in press.
(21) Erickson, J.; Kempf, D. Structure-based design of symmetric inhibitors of HIV-1 protease. *Arch. Virol.*, in press.
(22) Erickson, J. W. Design and structure of symmetry-based inhibitors of HIV-1 protease. *Perspect. Drug Dis. Design* **1993**, *1*, 109–128.
(23) Kempf, D. J.; Norbeck, D. W.; Codacovi, L.; Wang, X. C.; Kohlbrenner, W. F.; Wideburg, N. E.; Saldivar, A.; Craig-Kennard, A.; Vasavanonda, S.; Clement, J. J.; Erickson, J. Structure-based inhibitors of HIV protease. In *Recent Advances in the Chemistry of Anti-infective Agents*; Bentley, P. H., Ponsford, R., Eds.; Royal Society of Chemistry: Cambridge, 1993; pp 297–313.
(24) Wlodawer, A.; Erickson, J. W. Structure-based inhibitors of HIV-1 protease. *Annu. Rev. Biochem.* **1993**, *62*, 543–585.
(25) Appelt, K. Crystal structures of HIV-1 protease-inhibitor complexes. *Perspect. Drug Dis. Design* **1993**, *1*, 23–48.
(26) Toh, H.; Ono, M.; Saigo, K.; Miyata, T. Retroviral protease-like sequence in the yeast transposon Tyl. *Nature* **1985**, *315*, 691–692.
(27) Katoh, I.; Yasunaga, T.; Ikawa, Y.; Yoshinaka, Y. Inhibition of retroviral protease activity by an aspartyl proteinase inhibitor. *Nature* **1987**, *329*, 654–656.
(28) Davies, D. R. The structure and function of the aspartic proteinases. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 189–215.
(29) Tang, J.; James, M. N. G.; Hsu, I. N.; Jenkins, J. A.; Blundell, T. L. Structural evidence for gene duplication in the evolution of the acid proteases. *Nature* **1978**, *271*, 618–621.
(30) Miller, M.; Jaskolski, M.; Rao, J. K. M.; Leis, J.; Wlodawer, A. Crystal structure of a retroviral protease proves relationship to aspartic protease family. *Nature* **1989**, *337*, 576–579.
(31) Navia, M. A.; Fitzgerald, P. M. D.; McKeever, B. M.; Leu, C. T.; Heimbach, J. C.; Herber, W. K.; Sigal, I. S.; Darke, P. L.; Springer, J. P. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **1989**, *337*, 615–620.
(32) Wlodawer, A.; Miller, M.; Jaskolski, M.; Sathyanarayana, B. K.; Baldwin, E.; Weber, I. T.; Selk, L. M.; Clawson, L.; Schneider, J.; Kent, S. B. M. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science* **1989**, *245*, 616–621.
(33) Lapatto, R.; Blundell, T.; Hennings, A.; Overington, J.; Wilderspin, A.; Wood, S.; Merson, J. R.; Whittle, P. J.; Danley, D. E.; Geoghegan, K. F.; Hawrylik, S. J.; Lee, S. E.; Scheld, K. G.; Hobart, P. M. X-ray analysis of HIV-1 proteinase at 2.7 Å resolution confirms structural homology among retroviral enzymes. *Nature* **1989**, *342*, 299–302.
(34) Spinelli, S.; Liu, Q. Z.; Alzari, P. M.; Hirel, P. H.; Poljak, R. J. The three-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU. *Biochimie* **1991**, *73*, 1391–1396.

(35) York, D. M.; Darden, T. A.; Pedersen, L. G.; Anderson, M. W. Molecular dynamics simulation of HIV-1 protease in a crystalline environment and in solution. *Biochemistry* **1993**, *32*, 1443–1453.
(36) Rao, J. K. M.; Erickson, J. W.; Wlodawer, A. Structural and evolutionary relationships between retroviral and eucaryotic aspartic proteinases. *Biochemistry* **1991**, *30*, 4663–4671.
(37) Greenlee, W. J. Renin inhibitors. *Med. Res. Rev.* **1990**, *10*, 173–236.
(38) Meek, T. D. Inhibitors of HIV-1 protease. *J. Enz. Inhib.* **1992**, *6*, 65–98.
(39) Huff, J. R. HIV protease: a novel chemotherapeutic target for AIDS. *J. Med. Chem.* **1991**, *34*, 2305–2314.
(40) Norbeck, D. W.; Kempf, D. J. HIV protease inhibitors. *Annu. Rep. Med. Chem.* **1991**, *26*, 141–150.
(41) Tomasselli, A. G.; Howe, W. J.; Sawyer, T. K.; Wlodawer, A.; Heinrikson, R. L. The complexities of AIDS: an assessment of the HIV protease as a therapeutic target. *Chim. Oggi.* **1991**, *9*, 6–27.
(42) Plattner, J. J.; Norbeck, D. W. Obstacles to drug development from peptide leads. In *Drug Discovery Technology*; Clark, C. R., Moos, W. H., Eds.; Horwood: Chichester, UK, 1990; pp 92–126.
(43) Erickson, J.; Neidhart, D. J.; VanDrie, J.; Kempf, D. J.; Wang, X. C.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse, J. W.; Turon, M.; Wideburg, N.; Kohlbrenner, W. E.; Simmer, R.; Helfrich, R.; Paul, D. A.; Knigge, M. Design, activity, and 2.8 Å crystal structure of a C_2 symmetric inhibitor complexed to HIV-1 protease. *Science* **1990**, *249*, 527–533.
(44) Kempf, D. J.; Norbeck, D. W.; Codacovi, L.; Wang, X. C.; Kohlbrenner, W. E.; Wideburg, N. E.; Paul, D. A.; Knigge, M. F.; Vasavanonda, S.; Craig-Kennard, A.; Saldivar, A.; Rosenbrook, J. W.; Clement, J. J.; Plattner, J. J.; Erickson, J. Structure-based, C_2 symmetric inhibitors of HIV protease. *J. Med. Chem.* **1990**, *33*, 2687–2689.
(45) Rogers, E. F. The use and design of biligand enzyme inhibitors. *J. Theor. Biol.* **1984**, *109*, 471–474.
(46) Beddell, C. R.; Goodford, P. J. Symmetry in prostaglandins. *Prostaglandins* **1977**, *13*, 494–502.
(47) Pearl, L. H.; Taylor, W. R. A structural model for the retroviral proteases. *Nature* **1987**, *329*, 351–354.
(48) Suguna, K.; Padlan, E. A.; Smith, C. W.; Carlson, W. D.; Davies, D. R. Binding of a reduced peptide inhibitor to the aspartic proteinase from Rhizopus chinensis: implications for a mechanism of action. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7009–7013.
(49) Plattner, J. J.; Greer, J.; Fung, A. K.; Stein, H.; Kleinert, H. D.; Sham, H. L.; Smital, J. R.; Perun, T. J. Peptide analogues of angiotensinogen. Effect of peptide chain length on renin inhibition. *Biochem. Biophys. Res. Commun.* **1986**, *139*, 982–990.
(50) Matayoshi, E. D.; Wang, G. T.; Krafft, G. A.; Erickson, J. Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer. *Science* **1990**, *247*, 954–958.
(51) Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* **1989**, *246*, 1149–1152.
(52) Luly, J. R.; BaMaung, N.; Soderquist, J.; Fung, A. K. L.; Stein, H.; Kleinert, H. D.; Marcotte, P. A.; Egan, D.; Bopp, B.; Merits, I.; Bolis, G.; Greer, J.; Perun, T.; Plattner, J. J. Renin inhibitors. Dipeptide analogues of angiotensinogen utilizing a dihydroxyethylene transition-state mimic at the scissile bond to impart greater inhibitory potency. *J. Med. Chem.* **1988**, *31*, 2264–2276.
(53) Chenera, B.; Boehm, J. C.; Dreyer, G. B. Synthesis of C_2-symmetric and pseudosymmetric HIV-1 protease inhibitors from D-mannitol and D-arabitol. *Bioorg. Med. Chem. Lett.* **1991**, *4*, 219–222.
(54) Kempf, D. J.; Codacovi, L.; Wang, X. C.; Kohlbrenner, W. E.; Wideburg, N. E.; Saldivar, A.; Vasavanonda, S.; Marsh, K. C.; Bryant, P.; Sham, H. L.; Green, B. E.; Betebenner, D. A.; Erickson, J.; Norbeck, D. W. Symmetry-based inhibitors of HIV protease. Structure-activity studies of acylated 2,4-diamino-1,5-diphenyl-3-hydroxypentane and 2,5-diamino-1,6-diphenylhexane-3,4-diol. *J. Med. Chem.* **1993**, *36*, 320–330.
(55) Kempf, D. J.; Marsh, K.; Paul, D. A.; Knigge, M. F.; Norbeck, D. W.; Kohlbrenner, W. E.; Codacovi, L.; Vasavanonda, S.; Bryant, P.; Wang, X. C.; Wideburg, N. E.; Clement, J. J.; Plattner, J. J.; Erickson, J. Antiviral and pharmacokinetic properties of C_2 symmetric inhibitors of the human immunodeficiency virus type 1 protease. *Antimicrob. Agents Chemother.* **1991**, *35*, 2209–2214.
(56) Hosur, M. V.; Bhat, T. N.; Kempf, D. J.; Baldwin, E. T.; Liu, B.; Gulnik, S.; Wideburg, N. E.; Norbeck, D. W.; Appelt, K.; Erickson, J. W. Influence of stereochemistry on activity and binding modes for C_2 symmetry-based inhibitors of HIV-1 protease. *J. Am. Chem. Soc.*, in press.
(57) Dreyer, B. G.; Boehm, J. C.; Chenera, B.; DesJarlais, R. L.; Hassell, A. M.; Meek, T. D.; Tomaszek, T. A.; Lewis, M. A symmetric inhibitor binds HIV-1 protease asymmetrically. *Biochemistry* **1993**, *32*, 937–947.
(58) Norbeck, D.; Kempf, D.; Marsh, K.; Paul, D.; Knigge, M.; Vasavanonda, S.; Clement, J.; Kohlbrenner, W.; Sham, H.; Wideburg, N.; Plattner, J. Discovery of potent, orally bioavailable inhibitors of HIV protease. In *Discovery of Potent, Orally Bioavailable Inhibitors of HIV Protease*; VII International Conference on AIDS; Amsterdam, 1992; in press.
(59) Abdel-Meguid, S. S.; Zhao, B.; Murthy, K. H. M.; Winborne, E.; Choi, J. K.; DesJarlais, R. L.; Minnich, M. D.; Culp, J. S.; Debouck, C.; Tomaszek, T. A.; Meek, T. D.; Dreyer, G. B. Inhibition of human immunodeficiency virus-1 protease by a C_2-symmetric phosphinate. Synthesis and crystallographic analysis. *Biochemistry* **1993**, *32*, 7972–7980.
(60) Johnson, V. A.; Merrill, D. P.; Chou, T. C.; Hirsch, M. S. Human immunodeficiency virus type 1 (HIV-1) inhibitory interactions between protease inhibitor Ro 31-8959 and zidovudine, 2',3'-dideoxycytidine, or recombinant interferon-alpha A against zidovudine-sensitive or -resistant HIV-1 in vitro. *J. Infect. Dis.* **1992**, *166*, 1143–1146.
(61) Lam, P. Y.-S.; Eyermann, C. J.; Hodge, C. N.; Jadhav, P. K.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, N. Y.; Chang, C. H.; Weber, P. C.; Jackson, D. A.; Sharpe, T. R.; Erickson-Viitanen, S. De novo design and discovery of potent nonpeptide HIV protease inhibitors. Presented at the 205th American Chemical Society National Meeting; Denver, CO, 1993; No. 96.
(62) Maren, T. H. Carbonic anhydrase—general perspectives and advances in glaucoma research. *Drug. Dev. Res.* **1987**, *10*, 255–276.
(63) Maren, T. H. The links among biochemistry, physiology, and pharmacology in carbonic anhydrase mediated systems. In *Carbonic Anhydrase From Biochemistry and Genetics to Physiology and Clinical Medicine, International Workshop on Carbonic Anhydrase*; Botre, F., Gros, G., Storey, B. T., Eds.; VCH Publishers: New York, 1991; pp 186–207.
(64) Maren, T. H. Direct measurements of the rate constants of sulfonamides with carbonic anhydrase. *Mol. Pharm.* **1992**, *41*, 419–426.
(65) Eriksson, A. E.; Jones, T. A.; Liljas, A. Crystallographic studies of human carbonic anhydrase II. In *Progress in Inorganic Biochemistry and Biophysics, Zinc Enzymes*; Bertini, I., Luchinat, C., Maret, W., Zeppezauer, M., Eds.; Birkhauser Boston, Inc.: New York, 1986; Vol. 1, pp 317–328.
(66) Eriksson, A. E.; Jones, T. A.; Liljas, A. Refined structure of human carbonic anhydrase II at 2.0 Å resolution. *Proteins* **1988**, *4*, 274–282.
(67) Coleman, J. E. Chemical reactions of sulfonamides with carbonic anhydrase. *Annu. Rev. Pharmacol.* **1975**, *15*, 221–242.
(68) Lindskog, S. The structural basis of kinetic differences between carbonic anhydrase isoenzymes. In *Progress in Inorganic Biochemistry and Biophysics, Zinc Enzymes*; Bertini, I., Luchinat, C., Maret, W., Zeppezauer, M., Eds.; Birkhauser Boston, Inc.: New York, 1986; Vol. 1, pp 307–316.
(69) Silverman, D. N.; Lindskog, S. The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water. *Acc. Chem. Res.* **1988**, *21*, 30–36.
(70) Christianson, D. W. Structural biology of zinc. *Adv. Prot. Chem.* **1991**, *42*, 281–355.
(71) Baldwin J. J.; Ponticello, G. S.; Anderson, P. S.; Christy, M. E.; Murcko, M. A.; Randall, W. C.; Schwam, H.; Sugrue, M. F.; Springer, J. P.; Gautheron, P.; Grove, J.; Mallorga, P.; Viader, M.-P.; McKeever, B. M.; Navia, M. A. Thienothiopyran-2-sulfonamides: novel topically active carbonic anhydrase inhibitors for the treatment of glaucoma. *J. Med. Chem.* **1989**, *32*, 2510–2513.
(72) Sola, M.; Lledos, A.; Duran, M.; Bertran, J. Anion binding and pentacoordination in zinc(II) complexes. *Inorg. Chem.* **1991**, *30*, 2523–2527.
(73) Alexander, R. S.; Nair, S. K.; Christianson, D. W. Engineering the hydrophobic pocket of carbonic anhydrase II. *Biochemistry* **1991**, *30*, 11064–11072.
(74) Fierke, C. A.; Calderone, T. L.; Krebs, J. F. Functional consequences of engineering the hydrophobic pocket of carbonic anhydrase II. *Biochemistry* **1991**, *30*, 11054–11063.
(75) Nair, S. K.; Calderone, T. L.; Christianson, D. W.; Fierke, C. A. Altering the mouth of a hydrophobic pocket. Structure and kinetics of human carbonic anhydrase II mutants at residue Val-121. *J. Biol. Chem.* **1991**, *266*, 17320–17325.
(76) Krebs, J. F.; Fierke, C. A.; Alexander, R. S.; Christianson, D. W. Conformational mobility of His-64 in the Thr-200 → Ser mutant of human carbonic anhydrase II. *Biochemistry* **1991**, *30*, 9153–9160.
(77) Jewell, D. A.; Tu, C.; Paranawithana, S. R.; Tanhauser, S. M.; LoGrasso, P. V.; Laipis, P. J.; Silverman, D. N. Enhancement of the catalytic properties of human carbonic anhydrase II by site-directed mutagenesis. *Biochemistry* **1991**, *30*, 1484–1490.
(78) Behravan, G.; Jonsson, B.-H.; Lindskog, S. Fine tuning of the catalytic properties of carbonic anhydrase. Studies of a Thr200 → His variant of human isoenzyme II. *Eur. J. Biochem.* **1990**, *190*, 351–357.
(79) Vallee, B. L.; Auld, D. S. Active-site zinc ligands and activated H_2O of zinc enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 220–224.

(80) Liang, J.-Y.; Lipscomb, W. N. Binding of substrate CO_2 to the active site of human carbonic anhydrase II: a molecular dynamics study. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3675–3679.
(81) Liang, J.-Y.; Lipscomb, W. N. Application of quantum mechanical calculations to studies of enzymatic catalysis and inhibitor binding of human carbonic anhydrase II. In *Structure and Methods, DNA Protein Complexes and Proteins, Proceedings of the Sixth Conversation in the Discipline Biomolecular Stereodynamics*; Sarma, R. H., Sarma, M. H., Eds.; Adenine Press: Schenectady, 1990; Vol. 2, pp 181–194.
(82) Tu, C.; Silverman, D. N.; Forsman, C.; Jonsson, B.-H.; Lindskog, S. Role of histidine 64 in the catalytic mechanism of human carbonic anhydrase II studied with a site-specific mutant. *Biochemistry* **1989**, *28*, 7913–7918.
(83) Tu, C.; Paranawithana, S. R.; Jewell, D. A.; Tanhauser, S. M.; LoGrasso, P. V.; Wynns, G. C.; Laipis, P. J.; Silverman, D. N. Buffer enhancement of proton transfer in catalysis by human carbonic anhydrase II. *Biochemistry* **1990**, *29*, 6400–6405.
(84) Coleman, J. E. Current concepts of the mechanism of action of carbonic anhydrase. In *Proceedings in Life Sciences: Biophysics and Physiology of Carbon Dioxide*; Bauer, C., Gros, G., Bartels, H., Eds.; Springer-Verlag: New York, 1980; pp 133–150.
(85) Appelt, K.; Bacquet, R. J.; Bartlett, C. A.; Boothe, C. L. J.; Freer, S. T.; Fuhry, M. M.; Gehring, M. R.; Hermann, S. M.; Howland, E. F.; Janson, C. A.; Jones, T. R.; Kan, C.; Kathardekar, V.; Lewis, K. K.; Marzoni, G. P.; Matthews, D. A.; Mohr, C.; Moomaw, E. W.; Morse, C. A.; Oatley, S. J.; Ogden, R. C.; Reddy, M. R.; Reich, S. H.; Schoettlin, W. S.; Smith, W. W.; Varney, M. D.; Villafranca, J. E.; Ward, R. W.; Webber, S.; Webber, S. E.; Welsh, K. M.; White, J. Design of Enzyme Inhibitors Using Iterative Protein Crystallographic Analysis. *J. Med. Chem.* **1991**, *34*, 1925–1934.
(86) Varney, M. D.; Marzoni, G. P.; Palmer, C. L.; Deal, J. G.; Webber, S.; Welsh, K. M.; Bacquet, R. J.; Bartlett, C. A.; Morse, C. A.; Booth, C. L. J.; Hermann, S. M.; Howland, E. F.; Ward, R. W.; White, J. Crystal-Structure-Based Design and Synthesis of Benz[cd]indole-Containing Inhibitors of Thymidylate Synthase. *J. Med. Chem.* **1992**, *35*, 663–676.
(87) Goodford, P. J. A computational procedure for determining energetically favourable binding sites on biologically important molecules. *J. Med. Chem.* **1985**, *28*, 849–857.
(88) Bartlett, P. A.; Shea, G. T.; Telfer, S. J.; Waterman, S. CAVEAT: a program to facilitate the structure-derived design of biologically active molecules. *Spec. Pub. R. Soc. Chem.* **1989**, *78*, 182–196.
(89) Bohm, H.-J. LUDI: Rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 593–606.
(90) DesJarlais, R. L.; Sheridan, R. P.; Seibel, G. L.; Dixon, J. S.; Kuntz, I. D.; Venkataraghavan, R. Using Shape Complementarity as an Initial Screen in Designing Ligands for a Receptor Binding Site of Known Three-Dimensional Structure. *J. Med. Chem.* **1988**, *31*, 722–729.
(91) VanDrie, J. H.; Weininger, D.; Martin, Y. C. ALADDIN: an integrated tool for computer-assisted molecular design and pharmacophore recognition from geometric, steric, and substructure searching of three-dimensional molecular structures. *J. Comput.-Aided Mol. Des.* **1989**, *3*, 225–251.
(92) Freire, E. Structural Thermodynamics: Prediction of Protein Stability and Protein Binding Affinities. *Arch. Biochem. Biophys.* **1993**, *303*, 181–184.
(93) Lee, F. S.; Chu, Z. T.; Bolger, M. B. Calculations of antibody-antigen interactions: microscopic and semi-microscopic evaluation of the free energies of binding of phosphorylcholine analogs to McPC603. *Protein Eng.* **1992**, *5*, 215–228.
(94) Lybrand, T. P.; McCammon, J. A.; Wipff, G. Theoretical calculation of relative binding affinity in host-guest systems. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 833–835.
(95) Miyamoto, S.; Kollman, P. A. Absolute and relative binding free energy calculations of the interaction of biotin and its analogs with streptavidin using molecular dynamics/free energy perturbation approaches. *Protein* **1993**, *16*, 226–245.
(96) Murphy, K. P.; Xie, D.; Garcia, K. C.; Amzel, L. M.; Freire, E. Structural Energetics of Peptide Recognition: Angiotensin II/Antibody Binding. *Protein* **1993**, *15*, 113–120.
(97) Novotny, J. Protein Antigenicity: A thermodynamic approach. *Mol. Immunol.* **1991**, *28*, 201–207.
(98) Reddy, M. R.; Bacquet, R. J.; Zichi, D.; Matthews, D. A.; Weish, K. M.; Jones, T. R.; Freer, S. Calculation of solvation and binding free energy differences for folate-based inhibitors of the enzyme thymidylate synthase. *J. Am. Chem. Soc.* **1992**, *114*, 10117–10122.
(99) Williams, D. H.; Searle, M. S.; Mackay, J. P.; Gerhard, U.; Maplestone, R. A. Toward an estimation of binding constants in aqueous solution: Studies of associations of vascomycin group antibiotics. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1172–1178.
(100) Richards, F. M. The protein folding problem. *Sci. Am.* **1991**, *264* (1), 54–63.
(101) Creighton, T. E. The protein folding problem. *Science* **1988**, *240*, 267, 344.
(102) Ring, C. S.; Cohen, F. E. Modeling protein structures: construction and their applications. *FASEB J.* **1993**, *7*, 783–790.
(103) Yue, K.; Dill, K. A. Inverse protein folding problem: designing polymer sequences. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4163–4167.
(104) Sander, C. Inverting the protein-folding problem. *Biochem. Soc. Symp.* **1990**, *57*, 25–33.
(105) Bruccoleri, R. E.; Karplus, M. Conformational sampling using high-temperature molecular dynamics. *Biopolymers* **1990**, *29*, 1847–1862.
(106) Higo, J.; Collura, V.; Garnier, J. Development of an extended simulated annealing method: application to the modeling of complementary determining regions of immunoglobulins. *Biopolymers* **1992**, *32*, 33–43.
(107) Martin, A. C.; Cheetham, J. C.; Rees, A. R. Modeling antibody hypervariable loops: a combined algorithm. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9268–9272.
(108) Thornton, J. M. Modelling antibody combining sites: a review. *Ciba Found. Symp.* **1991**, *159*, 55–69.
(109) Bruccoleri, R. E.; Karplus, M. Prediction of the folding of short polypeptide segments by uniform conformational sampling. *Biopolymers* **1987**, *26*, 137–168.
(110) Fine, R. M.; Wang, H.; Shenkin, P. S.; Yarmush, D. L.; Levinthal, C. Predicting antibody hypervariable loop conformations. II. Minimization and molecular dynamics studies of MCPC603 from many randomly generated loop conformations. *Proteins* **1986**, *1*, 342–362.
(111) Moult, J.; James, M. N. G. An algorithm for determining the conformation of polypeptide segments in proteins by systematic search. *Proteins* **1986**, *1*, 146–163.
(112) Shenkin, P. S.; Yarmush, D. L.; Fine, R. M.; Wang, H.; Levinthal, C. Predicting antibody hypervariable loop conformation. I. Ensembles of random conformations for ringlike structures. *Biopolymers* **1987**, *26*, 2053–2085.
(113) Head, G. T.; Stillinger, F. H.; Arrecis, J. A strategy for finding classes of minima on a hypersurface: implications for approaches to the protein folding problem. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 11076–11080.
(114) Rabow, A. A.; Scheraga, H. A. Lattice neural network minimization. Application of neural network optimization for locating the global-minimum conformations of proteins. *J. Mol. Biol.* **1993**, *232*, 1157–1168.
(115) Ripoll, D. R.; Scheraga, H. A. The multiple-minima problem in the conformational analysis of polypeptides. III. An electrostatically driven Monte Carlo method: tests on enkephalin. *J. Protein Chem.* **1989**, *8*, 263–287.
(116) Ripoll, D. R.; Piela, L.; Vasquez, M.; Scheraga, H. A. On the multiple-minima problem in the conformational analysis of polypeptides. V. Application of the self-consistent electrostatic field and the electrostatically driven Monte Carlo methods to bovine pancreatic trypsin inhibitor. *Proteins* **1991**, *10*, 188–198.