Molecular Recognition of Protein−Ligand Complexes: Applications to Drug Design

Robert E. Babine* and Steven L. Bender

Agouron Pharmaceuticals, Inc., 3565 General Atomics Court, San Diego, California 92121-1122

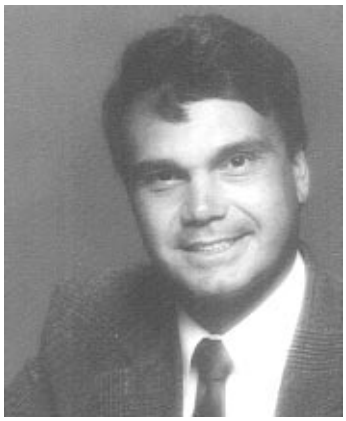
Received February 14, 1997 (Revised Manuscript Received May 19, 1997)

Contents

I. Introduction

Advances in molecular biology have had a dramatic impact on the drug discovery process.[1,2] The ability to produce recombinant proteins of biological significance has greatly impacted the screening process.[3] In addition, the ability to produce significant quantities of pure protein has facilitated the structure determination of many of these biologically relevant

S0009-2665(96)00370-6 CCC: $28.00



Robert E. Babine was born in Providence, RI, in 1955. He received his B.S. in Chemistry from the University of Rhode Island in 1977 and his Ph.D. in Organic Chemistry from Brown University in 1982. From 1982 until 1992 he was at Lederle Laboratories in Pearl River, NY. At that time he worked in the Medicinal Chemistry, Computational Chemistry, and Exploratory Medicinal Chemistry groups. His research involved β-lactam antibiotics, enediyne antitumor agents, and HIV protease inhibitors. From 1993 to present he has been a Senior Scientist in the Medicinal Chemistry group at Agouron Pharmaceuticals. His research efforts have been in the immunophilin area and in viral proteases. His research interest is structure-based design of enzyme inhibitors.

Steven L. Bender was born in Fairbury, IL, in 1958. After receiving a B.S. degree in Chemistry from the University of Illinois at Urbana–Champaign in 1981, he pursued graduate work first at the California Institute of Technology and then at Harvard University, where he received a Ph.D. degree in Organic Chemistry in 1986. In his graduate research with Professor David A. Evans, he completed a total synthesis of the polyether antibiotic X-206. From 1986 to 1988, he was a NIH postdoctoral fellow in the laboratory of Professor Jeremy Knowles at Harvard, where he studied the mechanism of the enzyme dehydroquinate synthase. From 1988 until 1992, he was an Assistant Professor in the Department of Chemistry at the University of California, Irvine, where he engaged in organic synthesis directed toward targets of biological interest. After a brief time at Ligand Pharmaceuticals from 1992 to 1993, he joined Agouron Pharmaceuticals, Inc., where he is now a Group Leader in the Medicinal Chemistry Department. His research efforts are focused on structure-based design of enzyme inhibitors.

targets. The availability of protein structures provides another tool in the drug design process.[1]

In an ideal situation, quantitative computational methods would greatly impact the protein structure-based design process. However, to date no methods are reliable enough, relative to qualitative human design, to be used in a prospective manner. A recent review has appeared which discusses the use of computational methods to assist in the drug design process and to evaluate protein−ligand complexes.[4]

The process of forming a complex between a small molecule ligand and a protein is a complex equilibrium process (Figure I*1). A solvated ligand likely exists as an equilibrium mixture of several conformers, likewise the solvated protein also exists as several conformers in equilibrium. To form a complex the solvent molecules that occupy the binding site[5,6] are displaced by the ligand to produce a solvated complex. Tight complexes will result when the protein−ligand interactions are significantly stronger than the interactions of the protein alone and ligand alone with solvent. From an entropic viewpoint, many aspects of complex formation are unfavorable since it results in the loss of conformational degrees of freedom for the both the protein and the ligand; in addition, three rotational and three translational degrees of freedom are lost upon complex formation.[7] Thus, highly favorable enthalpic contacts between the protein and the ligand must compensate for the entropic loss. As noted in Table I*1, small variations in complex stability measured in kilocalories per mole (kcal mol^{-1}) translate into significant differences in ligand affinity for a protein. A computational method with an accuracy of $\pm$3 kcal mol^{-1} would have errors in predicting affinities of more than 2 orders of magnitude. To date, no computational methods can reliably simulate the energetics of this equilibrium with acceptable precision.[8] Thus, the state-of-the-art design process relies, in large part, on a qualitative understanding of the molecular recognition of protein−ligand complexes based largely upon analogies to other systems. Our knowledge of these molecular recognition processes is enhanced by detailed analyses of structures of various protein−ligand complexes. This structural information is usually obtained from X-ray crystal

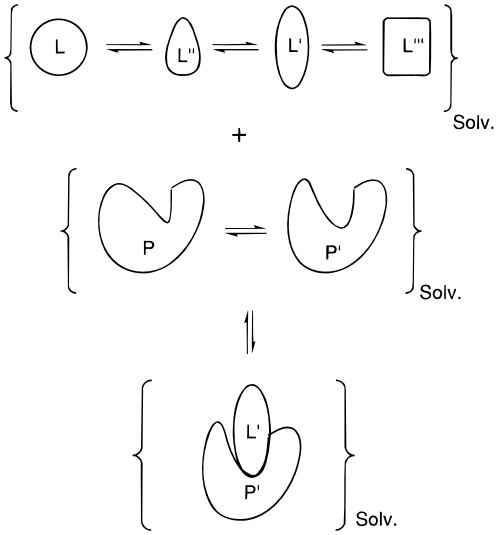


Figure I*1. Cartoon of protein−ligand complexation.

Table I*1. Correlation of Relative Protein−Ligand Affinities with Relative Stabilities of the Complexes at 300 K

rel K_d	ΔG (kcal mol^{-1})
5	0.96
10	1.37
29	2.00
100	2.73
156	3.00
840	4.00
4526	5.00

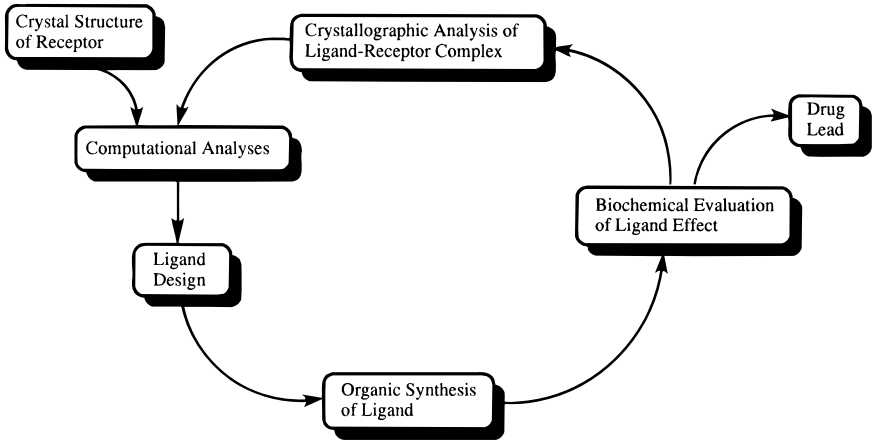


Figure I*2. The iterative protein structure-based design process: A system for the discovery and elaboration of lead compounds.

structures of the protein−ligand complex. Protein crystallography does not determine these structures to atomic resolution; therefore, there is some uncertainty in the exact position of each atom. The average errors in the atomic positions depend upon the diffracting quality of the protein crystal. For a resolution of about 2.4 Å, atom positions have been estimated to about 0.4 Å in space.[9−11] For structures determined with a resolution of less than 2.0 Å electron density is typically well defined and atomic positions are generally reliable and many water molecules are well resolved. For structures determined with a resolution of greater than 2.8 Å there is usually much uncertainty in the exact position of individual atoms, in addition, water molecules are not well defined in the electron density. As a general rule of thumb, the positional error of atoms amounts to approximately $^1/_6$ of the resolution.[4] X-ray diffraction of protein crystals does not allow for the determination of hydrogen atom positions.

Given a high resolution structure of a protein, new ligands can be designed completely *de novo*, in which in the ligand is designed purely on considerations of protein structure, or the design can be based on the modification of known ligands for the target protein. To date, the second method is the more common of the two.

The utilization of iterative design processes has greatly advanced the use of protein structure for the design of drug candidates (Figure I*2).[12,13] A crucial component of the iterative process involves the structure determination of complexes of newly prepared ligands with the protein of interest. By this process, structure and energy (i.e., experimental binding affinity) can be correlated, which greatly facilitates the optimization process. Our knowledge of protein−ligand molecular recognition processes has been enhanced by the application of these iterative cycles to many structurally diverse protein−ligand complexes. It has been written that "Ironically, detailed structures of these receptor−ligand complexes are sometimes easier to obtain than the structures of designed, synthetic host−guest complexes that are intended to model the more complex natural systems." [14]

To date, there are no comprehensive reviews which discuss the interactions of small molecule ligands with protein receptors. This review will fill this gap in the review literature. This review will cover the literature of the structures of complexes between small molecule ligands and several enzymes which are relevant as drug targets. Protein targets which are not enzymes are beyond the scope of this review. This review covers about 20 years of the literature ending in September 1996; however, most of the relevant work has been published within the last five years. Whenever possible the relevant Brookhaven Protein Data Bank (PDB)[15] entry has been given. The reader is encouraged to view these structures in three dimensions using a molecular graphics system. A major goal of this review is to enhance the qualitative understanding of protein−ligand complexes. Many examples of different ligands binding to the same protein will be provided. This should provide many analogies which should assist in the structure-based design process. While several classes of proteins will be reviewed, enzymes which perform chemistry on protein and polypeptide substrates will be discussed in greater detail. Many of these examples, in addition to illustrating protein−ligand interactions, also demonstrate the structural basis of peptide mimicry.

A major portion of this review provides an in-depth discussion of protease inhibitors. Other than reflecting biases of the authors this has several pedagogical advantages: (1) Four different mechanistic classes of enzymes that catalyze the same reaction, selective amide hydrolysis within a protein or polypeptide framework, are discussed. Successful strategies for inhibiting these enzymes vary from class to class and reflect the mechanistic and structural differences of the enzymes. (2) These enzymes have evolved to recognize proteins and polypeptides. While inhibition strategies based upon the preferred cleavage sites are now straightforward, they result in inhibitors which are peptide-like in structure. Peptides and peptidic compounds are usually unsuitable as drug candidates. This has resulted in the need for compounds which act as peptide mimetics. Many of the structures described will illustrate the structural basis for peptide mimicry. (3) There have been major research efforts to develop drug candidates as inhibitors of two different aspartic protease enzymes, renin and HIV-1 protease. In the search for renin inhibitors the structure of the target protein was not known

during most of this research. As a result, traditional medicinal chemistry approaches were used in an attempt to prepare drug candidates. To date there have been no renin inhibitors which are successful drug candidates. In contrast, during the search for inhibitors of HIV-1 protease, the structure of the protein was known and many groups have used an iterative structure-based design process. This has resulted in many structurally distinct classes of inhibitors. Several successful drug candidates have emerged from these studies and the use of HIV-1 protease structure in drug design represents one of the more impressive success stories in this emerging field.

II. Protease Enzymes: Introduction

Protease (proteinase) enzymes catalyze sequence selective amide hydrolysis. Amides are hydrolytically stable and protease enzymes catalyze an inherently slow reaction. The uncatalyzed rate of amide hydrolysis by water has been measured independently by three groups using three different experimental protocols. These three determinations have calculated the rate of amide hydrolysis (k_{non}) to be 3×10^{-9} s^{-1} at room temperature and neutral pH,[16] 1.25×10^{-9} s^{-1} at pH 9.0 and 37.0 °C,[17] and $(3.6-6.3) \times 10^{-11}$ s^{-1} at pH 6.8 and 25 °C.[18] For angiotensin-converting enzyme, a zinc metallo protease, the "catalytic proficiency",[19] defined as $(k_{cat}/K_m)/k_{non}$ (expressed in units of M^{-1}) is 5.3×10^{13}.[18] This value corresponds to 18.8 kcal mol^{-1} stabilization of the transition state for the amide hydrolysis reaction. The evolved ability of an enzyme to stabilize a transition state,[20] reflected by its catalytic proficiency, represents binding energy available to inhibitors which resemble highly activated reaction intermediates, i.e., transition-state analogs.[21] For the angiotensin-converting enzyme example, an inhibitor which could take full advantage of the ability of the enzyme to stabilize its transition state would be a 0.02 pM inhibitor.

Mechanistically, there are two broad classes of protease enzymes. One class employs an enzyme-bound activated water molecule as the nucleophile which attacks the amide carbonyl of the scissile bond. This water can either be activated by ligation to a zinc cation (the zinc metalloproteinases), or by binding in a small cleft defined by two aspartic acid residues (the aspartate proteases). These classes of enzymes catalyze direct amide hydrolysis by water. The second class of enzymes activate a nucleophilic atom of an amino acid residue to initiate amide hydrolysis. To date, the hydroxyl group of a serine or a threonine residue, and the thiol group of cysteine residue, have been identified as enzyme-activated nucleophiles. These classes of enzymes catalyze amide hydrolysis indirectly. The nucleophilic atom of the enzyme attacks the amide carbonyl of the scissile bond and breaks the C−N bond, to form an acyl enzyme intermediate (either an ester or a thioester). This acyl enzyme intermediate is then hydrolyzed by water to complete the hydrolysis process.

In addition to catalyzing amide hydrolysis, most proteases are capable of cleaving peptide bonds with sequence selectivity. This is usually accomplished by having an enzyme binding site(s) which is complimentary to one (or more) substrate residue(s). Standard nomenclature is used to designate substrate residues and their corresponding binding sites on the enzyme (Figure II*1).[22]

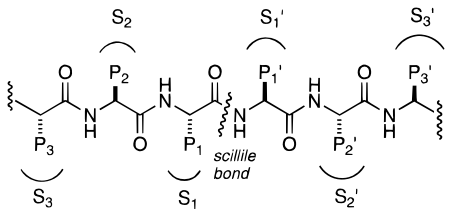


Figure II*1. Standard nomenclature, P_n, P_2, P_1, P_1', P_2', P_n', etc., are used to designate amino acid residues of a peptide substrate. The scissile peptide bond is P_1-P_1'. The corresponding binding sites on the protease enzyme are referred to as the S_n, S_2, S_1, S_1', S_2', S_n', etc. subsites.

III. Inhibition of Aspartic Proteases

A. Introduction

Two members of the aspartic protease class of enzymes, renin and HIV-1 protease, have been the subjects of extensive research efforts to discover therapeutically useful inhibitors. Renin inhibitors have potential as antihypertensive agents,[23] and HIV-1 protease inhibitors have been shown to be clinically useful in the control of AIDS.[24,25]

Aspartic proteases are endopeptidases. The primary sequence in this class has two different Asp-Thr-Gly sequences (for cathepsin D: PDB entry 1LYA Asp-33, Thr-34, Gly-35 and Asp-231, Thr-232, Gly-233). The apostructure of aspartic proteases (cathepsin D as a specific example)[26] shows these two chains running in opposite directions with a water molecule bound between the two aspartates. This water molecule is believed to be the nucleophile for the enzyme-catalyzed amide hydrolysis. (See Figure AP*1.) It has been concluded that the catalytically competent form of aspartic proteases is one in which one aspartate is protonated and the other is unprotonated (0,−1 protonation state).[27] The proposed hydrolysis mechanism involves general base-catalyzed addition of the water molecule to the scissile amide carbonyl to generate a tetrahedral intermediate, followed by proton transfers leading to a protonated nitrogen (general acid catalysis) and breakdown of that tetrahedral intermediate leading to the amide hydrolysis products.[28−31]

B. Inhibition of Aspartic Proteases by Pepstatin and Related Compounds

Pepstatin, **AP-1**, a natural product first isolated in 1970, was found to be a potent inhibitor of pepsin (K_i = 56 pM).[32] It was subsequently shown that pepstatin is a generic aspartic protease inhibitor and that pepstatin inhibition is characteristic of this enzyme class. Structure−activity data and crystal structures of pepstatin and its derivatives with various aspartic proteases have greatly enhanced our understanding of this enzyme class.[33] This information has proved extremely useful in developing potential drug candidates targeting these enzymes.

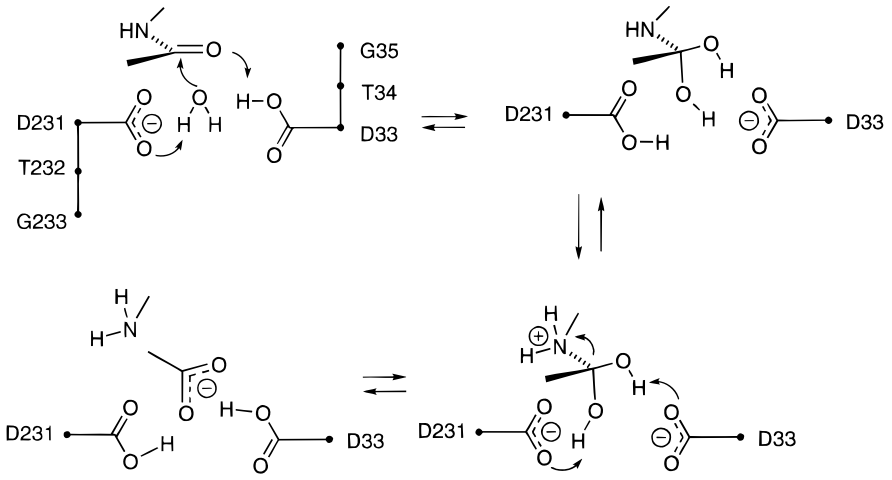


Figure AP*1. Proposed mechanism for aspartic protease amide hydrolysis. A nucleophilic water molecule held between the catalytic aspartates in the 0,−1 protonation state attacks the scissile amide carbonyl (general base catalysis) to generate a tetrahedral intermediate. Protonation of the amine in the tetrahedral intermediate (general acid catalysis) allows for the breakdown of the tetrahedral intermediate leading to products.

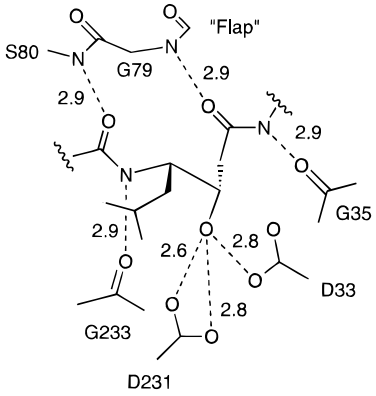
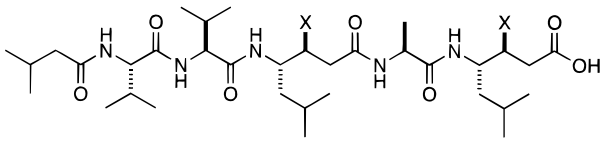


Figure AP*2. Schematic of interactions between pepstatin and pepsin. The central hydroxyl group interacts with both of the catalytic aspartates and displaces a structural water present in the apoenzyme. Protein backbone to inhibitor backbone hydrogen bonds orient and stabilize the inhibitor in the active site.

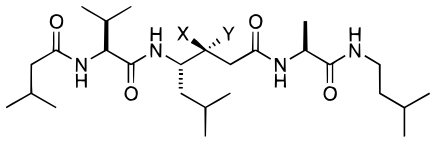
The structure of pepstatin bound to cathepsin D (PDB entry: 1LYB; see Figure AP*2)[34] shows several interesting features: (1) The hydroxyl group of the central statine ((3*S*,4*S*)-3-hydroxy-4-amino-6-methylheptanoic acid) unit has displaced the bound water between the two aspartate groups. This interaction is believed to be one of the most significant forces holding the complex together. (2) The central statine unit is acting as an isostere for a restricted conformation of a dipeptide unit. (3) All the groups of the peptide backbone from the P_4 to P_1' positions in pepstatin make complimentary hydrogen bonds with the enzyme. (4) The side chains from P_4 to P_3' all make VDW contact with the enzyme. In the case of cathepsin D they are favorable hydrophobic interactions. (5) The loop from His-77 to Gly-81 has undergone a conformational change in the complex, relative to the apoenzyme. This conformational change allows for formation of hydrogen bonds to the P_2 and P_1' carbonyl groups of pepstatin. In aspartic proteases this loop is often referred to as "the flap".

Much structure−activity relationship (SAR) information has been obtained concerning the role of the central statine unit of pepstatin and related analogs (see Table AP*1).[35−38] Comparison of the porcine

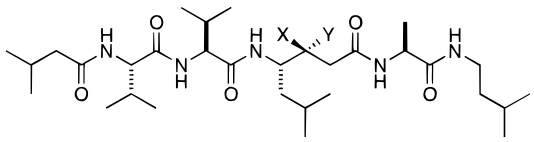
Table AP*1. Inhibition of Porcine Pepsin by Pepstatin and Analogs



		K_i (nM)
AP-1	X = OH	0.056
AP-2	X = H	210.



AP-3	X = OH, Y = H	1.1
AP-4	X = H, Y = OH	3000.
AP-5	X, Y = O	56.



AP-6	X = OH, Y = H	0.1
AP-7	X = H, Y = OH	100.
AP-8	X = Me, Y =OH	1.5
AP-9	X = OH, Y = Me	1200.

pepsin inhibitors **AP-1** and **AP-2** shows a 3750 fold loss of activity (4.9 kcal mol^{-1}) when the statine units are substituted with deoxy statine units. Compounds **AP-3** and **AP-4** show that the stereochemistry of the carbon bearing the hydroxyl group has a large effect (2727-fold; 4.7 kcal mol^{-1}) on protease inhibition. In contrast, comparison of **AP-1**, **AP-3**, and **AP-6** shows a relatively small loss in potency (less than 20-fold; 1.8 kcal mol^{-1}) as the ends of the inhibitor are truncated. These data suggest that the binding of

the central hydroxyl group between the catalytic aspartates plays a key role in complex stability. Ketone **AP-5** shows a 50-fold loss in potency (2.3 kcal mol^{-1}) compared to alcohol **AP-3**. ^{13}C NMR experiments have shown that the ketone is hydrated, by an enzyme-bound water molecule, within the active site of pepsin.[39] One interesting result is the reversed stereochemical preference for the alcohol groups in **AP-8** and **AP-9**. No structural information is available to explain this result.

For the apoenzyme, trapping the catalytic water molecule is entropically unfavorable, thus the enzyme-water enthalpic interaction must be greater than the entropic cost of trapping this water molecule. The hydroxyl group of pepstatin makes favorable enthalpic interactions with the aspartate groups. These interactions are similar to those made by the bound water, allowing the release of the bound water into bulk solvent. This water release process is highly favorable from an entropic standpoint. On the basis of standard entropies for anhydrous and hydrated inorganic salts, the standard entropy of ice at its freezing point and the standard entropy of liquid water, it has been estimated that the entropy cost of transferring a water molecule from bulk solvent into the interior of a protein lies between 0 and 7 cal mol^{-1} K^{-1}. The higher value only applies to those few water molecules that are most firmly bound.[40] The catalytic water molecule can be classified as being tightly bound to aspartic proteases. From these entropy estimates it has been suggested that the free energy cost of trapping a tightly bound water molecule is close to 2 kcal mol^{-1} at 300 K. Other estimates of the entropic cost of binding a water molecule to a protein, based upon measured adsorption and desorption isoterms of water vapor on solid proteins, suggest a larger value ranging from 3.6 to 6.0 kcal mol^{-1} at 300 K.[41] While the hydroxyl group makes favorable enthalpic interactions with the protein, these interactions can only be considered as replacements for the interactions made by the bound water molecule. Thus, the net enthalpic process of displacing this water is relatively small. The net energetics is a balance between entropy and enthalpy and is an example of enthalpy−entropy compensation.[42]

The most significant impact of this structure on inhibitor design is the realization of how much binding affinity can be gained by efficiently displacing the enzyme bound catalytic water molecule. The data in Table AP*1 suggests that making highly favorable interactions with the catalytic aspartates, thus displacing the catalytic water, can contribute about 5 kcal mol^{-1} to the stability of the complex.

Pepstatin is a slow binding inhibitor of many aspartic proteases. The $t_{1/2}$ for formation of the tightened pepstatin−pepsin complex is ~30 s. It has been proposed that the slow binding is due to the water extrusion process. Upon analysis of the kinetics of inhibition, it has been proposed that pepstatin resembles a collected substrate (i.e., peptide plus catalytic water) more than it does a transition state.[33]

In related studies, two fluorinated pepstatin analogs **AP-10** and **AP-11** were prepared. The difluoro statine derivative **AP-10** was found to have a K_i = 10 nM against penicillopepsin, the difluoro statone

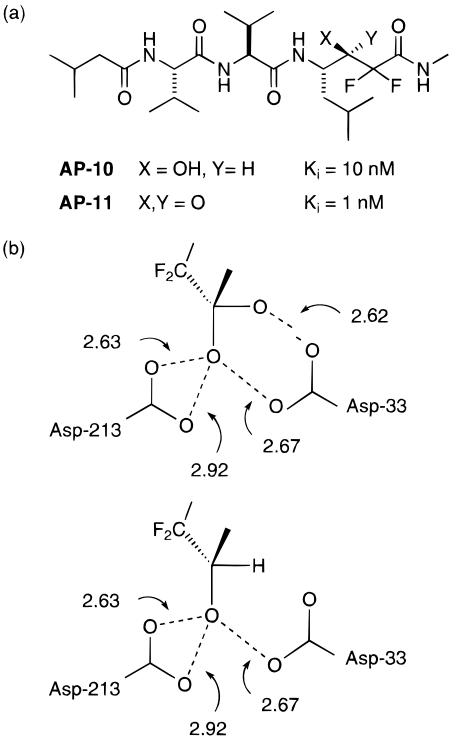


Figure AP*3. (a) Structures of various inhibitors and (b) comparison of the interactions between difluorostatone **AP-11** (bound as its hydrate) and difluorostatine **AP-10** with catalytic aspartates of penicillopepsin.

derivative **AP-11** was more potent with K_i = 1 nM. High-resolution structures for both compounds in complex with penicillopepsin have been reported (PDB entries: APV and APW).[31] Compound **AP-11** was observed to be bound to the enzyme as its hydrate. Both inhibitors adopt virtually identical conformations in their respective complexes that closely resemble the bound conformation observed for the central part of pepstatin bound to cathepsin D. For **AP-11**, one hydroxyl group is bound between the two catalytic aspartates while the second hydroxyl group interacts only with Asp-33. (See Figure AP*3.) Difluorostatone **AP-10** binds identically except that the hydrogen bond from the *pro-S* hydroxyl group is absent. The more potent inhibition of **AP-11** is presumably due to the presence of the additional hydrogen bond.

One interesting feature of aspartic protease inhibition is that the difluoro ketone **AP-11** is 10 times more potent than difluoro alcohol **AP-10**, while ketone **AP-5** is 50-fold less potent than the analogous alcohol **AP-3**. The observation that a difluoro ketone is a more potent inhibitor than a normal ketone is likely related to ease of hydration. For difluoro ketone **AP-11**, an ^{19}F NMR experiment in 80% DMSO/20% H_2O showed resonances only for the hydrated form.[31] For ^{13}C-labeled ketone **AP-5**, the hydrated ketone was not detected in a ^{13}C NMR experiment in buffer. It is likely that the inhibitory species for the difluoro ketone is its hydrate. Thus, the hydrate binds to the enzyme and displaces the catalytic water molecule. For ketone **AP-5**, the ketone binds to the enzyme and is hydrated in the active site by the catalytic water molecule.[39] The less potent inhibition of a ketone is likely due to strain introduced by an unfavorable hydration process.

Phosphinic acid analogs of pepstatin have been reported to be potent, slow-binding inhibitors of aspartic proteases.[43,44] Compound **AP-12** was determined to inhibit pepsin with a $K_i = 3.0$ nM. This compound is nearly equipotent with the corresponding pepstatin analog **AP-3**. This K_i represents the initial K_d of the enzyme−inhibitor complex; in a slow step ($t_{1/2} = 115$ min), the complex forms a much tighter complex with a K_i estimated at less than 70 pM. The related analog **AP-13** was determined to be a potent inhibitor of both pepsin and penicillopepsin with K_i's of 207 and 22 nM, respectively, at pH = 3.5. Penicillopepsin inhibition by this compound was determined to have a pH dependence. At pH = 4.5 it was a less potent inhibitor with a $K_i = 107$ nM. A crystal structure of the complex between **AP-13** and penicillopepsin has been reported (PDB entry: 1PPK).[45] A related phosphonate compound **AP-14** has been reported and has K_i's of 3.7 and 2.8 nM at pH = 3.5 against pepsin and penicillopepsin, respectively. Unlike **AP-12**, **AP-14** does not exhibit a slow-binding, two-step mechanism of inhibition. A close analog of **AP-14** also displays pH dependence of inhibition. A crystal structure of the complex between **AP-14** and penicillopepsin has also been reported (PDB entry: 1PPL).[45]

The structure of the complex between phosphinate **AP-13** and penicillopepsin is very similar to the complex between difluoro ketone **AP-11** and penicillopepsin. One of the phosphinate oxygens is bound between the two catalytic aspartates, while the second oxygen interacts only with Asp-33. However, there are rather subtle differences in the two complexes. For **AP-13** there is a very short oxygen−oxygen contact (2.39 Å) between the inhibitor and the Asp-33. (See Figure AP*4b.) Phosphonate **AP-14** shows nearly identical interactions between the phosphonate oxygens and the catalytic aspartates as does **AP-13**. The differences between the phosphorous inhibitors (**AP-13** and **AP-14**) and the hydrated difluoro ketone (**AP-11**) have been discussed in terms of different hydrogen bonding arrangements between these different classes of inhibitors.[45]

The pH dependence on inhibition can be explained by the experimentally observed close contact between the phosphinate oxygens of **AP-13** and the catalytic aspartates. The observed close contacts between the phosphinate and carboxylates would seem to be most consistent with the three functional groups sharing a −1 charge, and inconsistent with these three groups sharing a −2 charge. The pH dependence of inhibition is consistent with allowing either a protonated phosphinate to interact with an active site in the −1,0 protonation state (one aspartate deprotonated and the other protonated), or a charged phosphinate interacting with an active site in the 0,0 state (both aspartates protonated). Either of these two situations will be more likely at lower pH, consistent with the observed pH dependence of inhibition. One limitation of protein crystallography is the inability to determine the positions of hydrogen atoms and determine the hydrogen bonding arrangement. The same argument holds for phosphonate **AP-14**. Speculation on active protonation states for **AP-13** and **AP-14** and the differences in protonation state for **AP-11** has been published.[45]

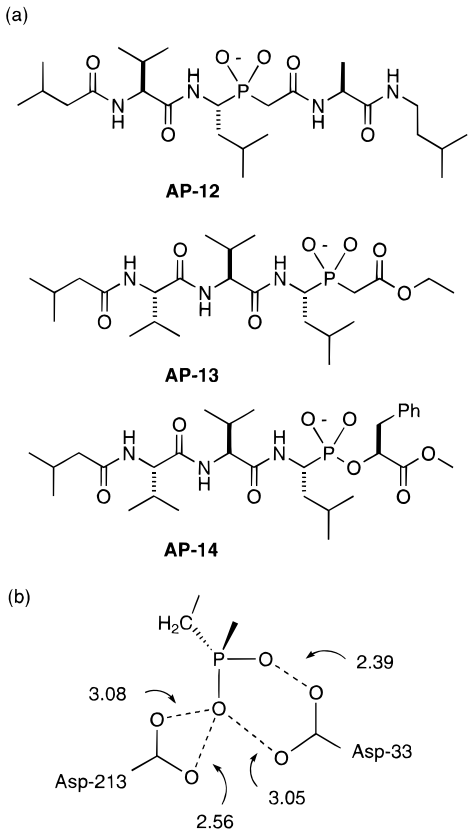


Figure AP*4. (a) Structures of various inhibitors and (b) interactions between the phosphinate group of **AP-13** and the catalytic aspartates of penicillopepsin. Notice the close O−O protein−inhibitor contacts, it is likely that these three groups share a total charge of −1.

A series of 11 phosphonate peptide analogs, RO_2C-Xaa-Yaa-Phe-{PO_2^--O}-Phe-O-(3-(4-pyridyl)propyl ester), related to **AP-14**, were synthesized and evaluated as inhibitors of porcine pepsin.[46] For eight of these inhibitors, with Gly or Ala at the P_2 position, the K_i values correlate with the K_m/k_{cat} values of the corresponding substrates, demonstrating that these analogs mimic the transition state for peptide bond hydrolysis. While **AP-14** has not experimentally been shown to be a transition state analog of penicillopepsin, this complex likely represents the structure closest to a "transition-state structure" that has been reported to date.

Of the various natural products for which structures have been obtained for complexes with their biological receptor, those between pepstatin and aspartic proteases have undoubtedly had the most impact on drug design strategies.[47] The following section on HIV-1 protease inhibitors will expand on the importance of this contribution.

C. Inhibition of HIV-1 Protease

1. Introduction

In the last several years, a major research effort in several laboratories has been the search for clinically useful inhibitors of HIV-1 protease (HIVp). During these studies many structures of diverse inhibitors in complex with HIVp have been reported.

HIVp is an aspartic protease and the first reported inhibitor was the natural product pepstatin.[48] The structure between HIVp and acetyl-pepstatin has been reported (PDB entry: 5HVP).[49]

For many years, several research groups were involved in searching for orally bioavailable inhibitors of another aspartic protease, renin, for the control of hypertension.[50] To date, no clinically useful renin inhibitors have been developed. A major stumbling block in the development of renin inhibitors was the inability to prepare compounds with acceptable oral bioavailability. Most of the renin inhibitors reported have been peptidomimetics that retain significant peptidic character. Most of the design principles for renin inhibition were derived from the classical studies of the inhibition of fungal aspartic proteases by pepstatin and related compounds. During most of the renin inhibitor research the 3D structure of the renin enzyme was not known.[51]

In contrast to renin, the 3D structure of HIVp was available soon after the identification of HIVp as a target for therapeutic intervention.[52,53] The initial starting point for the design of HIVp inhibitors were the principles of aspartic protease inhibition developed during the search for renin inhibitors. However, as HIVp research progressed, more and more cocrystal structures became available. Ultimately, from random screening, trial and error SAR studies, and structure based design strategies, clinically useful nonpeptidic HIVp inhibitors have resulted. Since the HIVp inhibitor field has been reviewed previously,[54−57] this section will focus primarily the following: (1) The diversity of HIVp inhibitors and a discussion of their ligand−protein interactions. (2) An extension of the principles of aspartic protease inhibition. This will be achieved by examining the structures of varied HIVp inhibitors and their interactions with the catalytic aspartates. (3) Discussion of novel peptidomimetic groups and their interactions with the protein. (4) A discussion of emerging principles of how to decrease peptidic character and inhibitor size while retaining potent enzyme inhibition.

Since the structure and function of HIVp has been previously reviewed it will not be extensively discussed here. HIVp is a C_2 symmetrical homodimer with each monomer having 99 residues. The enzyme's C_2 axis lies between and perpendicular to the catalytic aspartates (Asp-25 and Asp-25′) in the active site. Because of the enzyme's symmetric nature, HIVp has not evolved structurally distinct substrate C-terminal and N-terminal domains. Using standard nomenclature,[22] the S_1 and S_1' (S_2 and S_2', etc.) subsites are structurally identical. The two equivalent S_1 subsites are very hydrophobic, the S_2 subsites are mostly hydrophobic, except for the Asp-29, Asp-30, Asp-30, and Asp-30′. The S_3 subsites are adjacent to the S_1 subsites and are mostly hydrophobic with the exception of Arg-8 and Arg-8′. (See Figure AP*5.)

2. Hydroxyethylene- and Hydroxyethylamine-Based Inhibitors

The substrate-based inhibitor **U-85548e** contains a hydroxyethylene isostere replacement at the scissile bond that is believed to mimic the tetrahedral transition state of the proteolytic reaction. This compound has a K_i < 1 nM against HIVp. The classic studies on aspartic protease inhibition by pepstatin clearly laid the groundwork for the design of this potent inhibitor. A structure of the complex between **U-85548e** and HIVp was reported by the Upjohn and NCI groups (PDB entry: 8HVP; see Figure AP*6).[58] A detailed discussion of this structure is warranted since it is a large inhibitor, spanning P_5 to P_3', and it illustrates many of the protein−ligand contacts available for inhibitors.

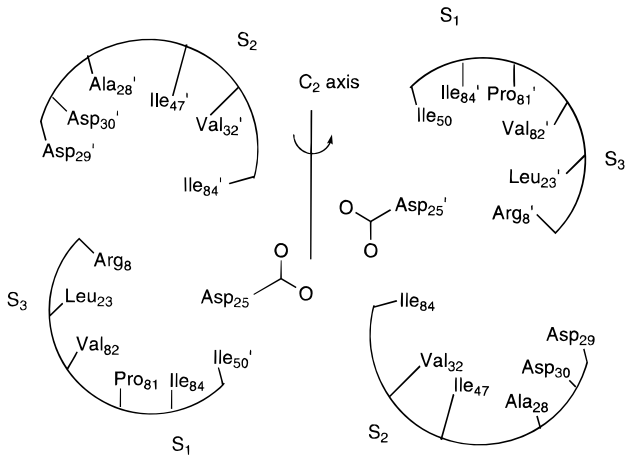


Figure AP*5. Amino acid residues which form binding sites of HIVp.

The inhibitor lies within the large HIVp active site in an extended conformation and makes protein−ligand contacts from P_4 to P_3'. The hydroxyl group of hydroxyethylene moiety lies between the two catalytic aspartates (D25 and D25′) and makes favorable electrostatic contacts in much the same manner that pepstatin does with cathepsin D. Like the case of pepstatin, this interaction is believed to be one of the most significant forces holding the complex together.

The peptide backbone of the inhibitor makes hydrogen bond contacts with the main chain of the protein. The amide linkage between P_2 and P_3 makes a strong hydrogen bond between its NH and the carbonyl oxygen of Gly-48 and a longer hydrogen bond between its carbonyl group and the backbone NH of Asp-29. At the symmetry-related site, the amide linkage between P_2' and P_3' makes a strong hydrogen bond between its NH and the carbonyl oxygen of Gly-48′ and a longer hydrogen bond between its carbonyl group and the backbone NH of Asp-29′. For the amide linkage between P_3 and P_4 the carbonyl group makes a strong hydrogen bond with the backbone NH of Gly-48 and the NH group makes an electrostatic interaction with the side chain of Asp-29. At the symmetry-related site, the carbonyl group of the amide linkage between P_3' and P_4' makes a strong hydrogen bond with the backbone NH of Gly-48′ and the NH group makes an electrostatic interaction with the side chain of Asp-29′. Long hydrogen bonds are also formed between the amide NH linking the P_1 and P_2 groups and the backbone carbonyl of Gly-27, as well as at the symmetry related Gly-27′ with the amide NH between the P_1' and P_2' groups. At the center of the inhibitor an interesting hydrogen-

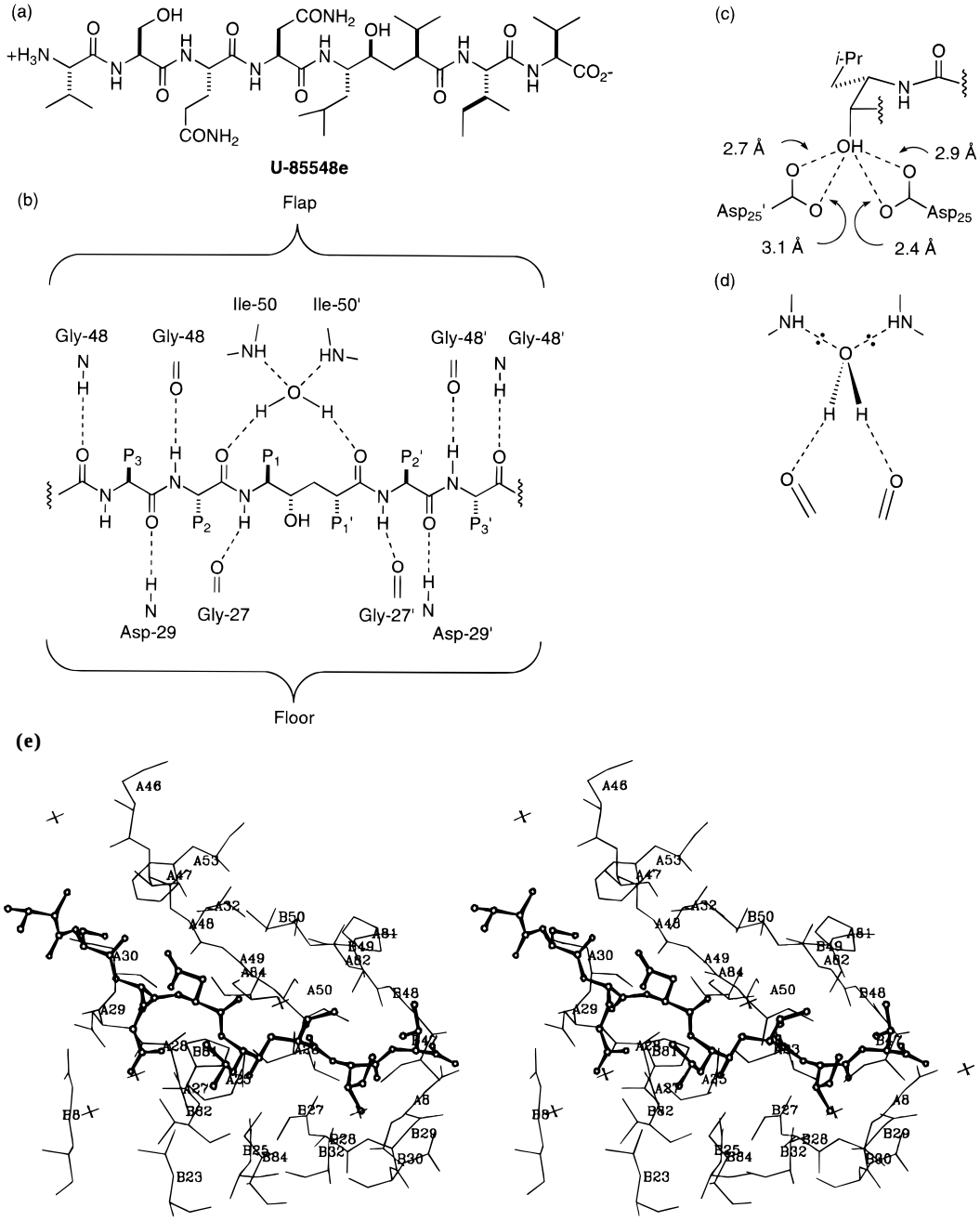


Figure AP*6. (a) Structure of peptidyl inhibitor **U-85548e**. (b) Summary of protein−ligand hydrogen bonds between **U-88548e** and HIVp. The flap water makes hydrogen bonds between both the protein and the inhibitor. (c) Interactions of central hydroxyl group of **U-85548e** with the catalytic aspartates of HIVp. (d) The flap water, present in most HIVp−inhibitor complexes, is tetrahedrally coordinated. (e) Stereoview of the HIVp **U-85548e** complex.

bond arrangement is observed. The carbonyl group between P_1 and P_2 and the carbonyl group between P_1' and P_2' both make hydrogen bonds to the same water molecule. This water molecule also makes two hydrogen bonds to the protein, one with the backbone NH of Ile-50 and another with the backbone NH of Ile-50′. This water molecule is completely satisfied by making hydrogen bonds to both the inhibitor and the protein with approximately tetrahedral geometry. The presence of this water molecule has been observed for the majority of inhibitor−HIVp complexes reported to date and will be referred to as the "flap" water throughout the remainder of this review. The presence of the flap water introduces a deviation from C_2 symmetry in the protein (residues Gly-49, Ile-50, Gly-51; and Gly-49′, Ile-50′, Gly-51′). The hydrogen bonding pattern observed for **U-85548e** is typical for many peptidic type inhibitors and is illustrated in Figure AP*6b.

Two structural waters are seen in the floor of the enzyme active site, between the S_2 and S_3 subsites. The more deeply buried of these waters interacts with the backbone carbonyl of Thr-26 and the side chain of Arg-87. The other water interacts with the backbone carbonyl of Gly-27, the anti lone pair of the carboxylate of Asp-29 and the carbonyl group of the P_3 Gln group of **U-85548e**. These two water molecules are 3.05 Å apart. At the symmetry-related site, only one water is present and it interacts with the backbone carbonyl of Gly-27′, the anti lone pair of the carboxylate of Asp-29′. These structural waters will be referred to as the "floor" waters. In

many of the HIVp inhibitor complexes seen, one or two floor waters are present in each of the symmetry-related sites and often interact with the bound inhibitor.

The P_1, P_1', P_2', and P_3' side chains make hydrophobic contacts with the protein. The P_3 glutamine side chain makes both VDW contacts with the protein and a hydrogen bond to one of the "floor" water molecules. The P_3 and P_1 groups of the inhibitor make VDW contacts with each other and generate a composite surface. This composite surface makes VDW contacts with the adjacent S_1 and S_3 subsites of the protein. The P_4 serine side chain makes a hydrogen bond to the side chain of Asp-30. The P_2 asparagine side of **U-85548e** does not make any hydrogen bonds to the protein and appears to only make VDW contacts with the protein. The P_2 and P_4 groups make VDW contacts with each other, this composite P_2 and P_4 surface makes VDW contacts with the adjacent S_2 and S_4 subsites of the protein. The P_1' and P_3' valine groups of **U-85548e** make VDW contacts with each other. This composite P_1' and P_3' surface makes VDW contacts with symmetry-related S_1 and S_3 subsites.

Previous review articles have been published which discuss, in detail, the interactions of peptide-like inhibitors with HIVp.[54−57]

The SmithKline group reported a structure−activity analysis using enzyme inhibition and X-ray crystallography (PDB entries: 1AAQ, 1HEF, 1HEG) for hydroxyethylene isostere inhibitors of HIVp.[59] This study involved 20 compounds and four crystal structures. From this study, a minimal inhibitor model for tight binding was derived in which the inhibitor spans P_2-P_2', making optimal VDW contacts with the protein. From the perspective of drug design, reduction of inhibitor molecular weight is desirable. These results demonstrate that while interactions between P_3 and S_3 subsites can contribute to inhibitor affinity, they are not required for potent inhibition. In addition, the central hydroxyl group interacts with the catalytic aspartates, the flap water is present and is tetrahedrally coordinated, and four protein ligand hydrogen bonds are present. Inhibitor **AP-15** (MW = 511) was reported to have a $K_i = 1.4$ nM and represents a minimal hydroxyethylene isostere inhibitor.

While peptide-derived inhibitors can be potent HIVp inhibitors they are typically not suitable drug candidates. In general, peptide-like compounds possess undesirable physical properties (i.e., poor solubility), vulnerability to degradative enzymes, rapid biliary clearance, and poor oral bioavailability. A useful drug candidate, in addition to being a potent enzyme inhibitor, must have desirable pharmacological properties. That is, an HIVp inhibitor must be able to penetrate cell membranes to be potent antivirals, and they should be orally bioavailable. This has led to a major research undertaking to design nonpeptidic inhibitors of HIVp.

To reduce the peptidic nature of hydroxyethylene inhibitors the SmithKline group has reported on using an imidazole ring as an amide bond replacement. Crystal structures for two of these compounds, **SB203386**, $K_i = 18.0$ nM (PDB entry: 1SBG),[60] and

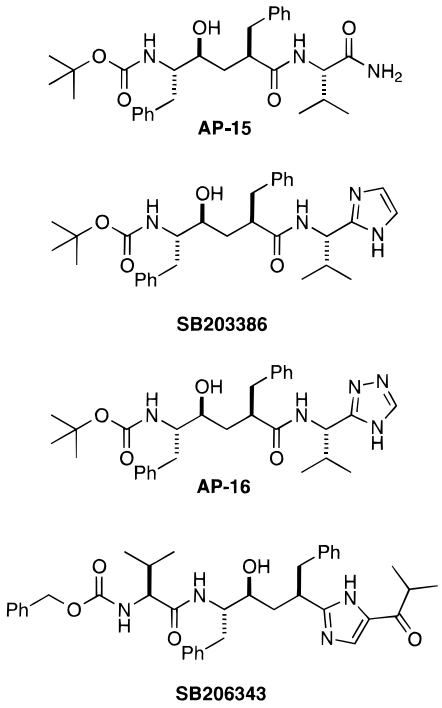


Figure AP*7. Structures of various inhibitors.

SB206343, $K_i = 0.6$ nM (PDB entry: 1HPS),[61] have been reported. (See Figure AP*7.)

Comparison of **SB203386** with **U-85548e** shows that the terminal imidazole make hydrogen bonds with the protein in the same fashion as does the amide group between P_2' and P_3'. At pH 6.0 it has a $K_i = 18.0$ nM and is an order of magnitude less potent than the minimal peptidic inhibitor **AP-15**. Compound **SB203386** exhibits a pH dependence on inhibition. It is a weaker inhibitor at lower pH. These data, and the crystal structures of their HIVp complexes, show that it is the neutral species that is a potent HIVp inhibitor. The observed structures show a strong hydrogen bond to the backbone carbonyl of Gly-48 and a weaker hydrogen bond to the backbone NH of Asp-29 and the lack of any strong interaction between the imidazole and the side chain of Asp-29. These structural data are consistent with the neutral species being the inhibitor. The related triazole analog **AP-16**, is a more potent inhibitor ($K_i = 4.2$ nM). Since the pK_a of a triazole, in general, is lower than that of an imidazole, the more potent inhibition by **AP-16**, which is presumably unprotonated at 6.0, suggests that the neutral forms of these compounds are the species that potently inhibits the enzyme.[60]

The imidazole ring in **SB206343** accepts a hydrogen bond from the flap water and donates a hydrogen bond to the carbonyl oxygen of Gly-27. The terminal carbonyl makes a hydrogen bond with the backbone NH of Asp-29 and the isopropyl group partly fills the S_2 subsite. The inhibition and structural data indicate that an imidazole, or a triazole ring, can be a useful amide bond surrogate.

The SmithKline group has also reported on attempts to prepare a constrained reduced-amide inhibitor of HIVp by the incorporation of a γ-turn mimetic for the P_2-P_1' residues into a model substrate. This compound, **SB203238**, was found to be

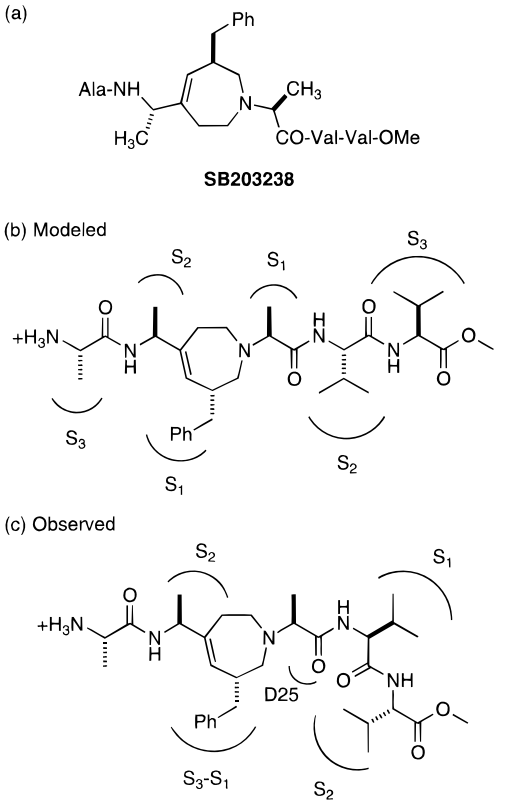


Figure AP*8. (a) Structure of constrained inhibitor **SB203238**. (b) Modeled interactions of **SB203238** with HIVp. The seven-membered ring was proposed as a γ-turn mimetic. (c) Observed interactions between **SB203238** and HIVp.

a competitive inhibitor of HIVp with a $K_i = 430$ nM.[62] A crystal structure of its complex with HIVp determined that it does not bind to the active site as modeled (PDB entry: 1HBV; see Figure AP*8).[63] Like other peptide analog inhibitors it is in an extended conformation. However, **SB203238** makes only two hydrogen bonds to the protein; one between the ester carbonyl and the backbone NH of Asp-29 and one between the amide NH of the γ-turn mimetic and the backbone carbonyl of Asp-29′. The flap water is present but is not tetrahedrally coordinated and makes only one hydrogen bond to the inhibitor (carbonyl group of the central valine). The amide carbonyl adjacent to the γ-turn mimetic interacts weakly with Asp-25. The benzyl group of **SB203238** makes hydrophobic contacts in one of the S_1 subsites and the adjacent S_3 subsite.

To dramatically reduce the peptide nature of inhibitors, the Agouron group[64−66] designed the simple compound **AG1132** ($K_i = 24$ μM). The crystal structure of its complex with HIVp shows the hydroxyl group interacting with the catalytic aspartates and hydrogen bond formation to the flap water from the two carbonyl groups. (See Figure AP*9.) For each of the benzamides, the ortho substituent favors the amide to be out of conjugation and allows hydrogen-bond formation with the flap water. The phenyl groups make VDW contacts with the S_1 subsites and the *tert*-butyl groups make VDW contacts with the S_2 subsites. The amide NH does not interact directly with the protein but rather interacts indirectly with the protein through ordered water molecules (the floor waters).

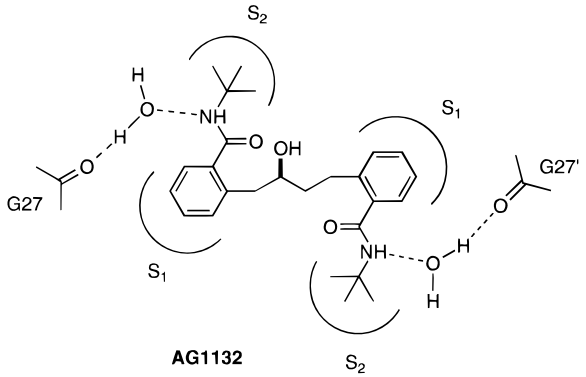


Figure AP*9. Hydrogen-bonding scheme observed for AG1132. The amide NH's interact with the protein through ordered water molecules. The "flap" water is present (not shown) and is tetrahedrally coordinated.

To interact with the S_2 and S_3 subsites more effectively, substitutions on the amide nitrogen were made. Compound **AG1157** ($K_i = 1.1$ μM) showed a 20-fold improvement in potency. A crystal structure of it bound to HIVp showed a number of unexpected features: (1) There was an inversion in binding mode; the phenyl groups were in the S_2 subsites and the *tert*-butyl groups were in the S_1 subsites. (2) This inversion in binding mode resulted in the enantiomer of the modeled compound being bound to the protein. This presumably allows simultaneous optimal hydrophobic interactions and allows for the hydroxyl to be between the aspartates with minimal strain in the linking chain. (3) The hydrogen bonds to the flap water occur with less than optimal geometry. The flap water is usually satisfied by hydrogen bonds with near tetrahedral geometry; this is not possible with **AG1157** and related compounds. (4) The hydroxyl group in the hydroxyethyl side chains have displaced structural (floor) waters which were present in the **AG1132** structure. (See Figure AP*10.)

The Agouron group then used the information of the novel binding mode of **AG1157** to optimize its interactions with HIVp. This was done by increasing the size of lipophilic groups and complimenting the morphology of the active site. This study resulted in the design of the potent, orally bioavailable, nonpeptidic, HIVp inhibitor **AG1284** ($K_i = 8$ nM). This study has shown that careful optimization of protein−ligand VDW contacts can contribute about 3.7 kcal mol^{-1}.

A similar binding mode inversion of P_1' and P_2' was observed by the Searle group for inhibitor **AP-17** that contained the tertiary urea functionality.[67] (See Figure AP*11.)

Another peptidyl inhibitor, **JG-365** (Figure AP*12), contains a hydroxyethylamine isostere replacement for the Phe−Pro cleavage site. This compound has a $K_i = 0.24$ nM against HIVp. The structure of this complex was reported by the University of Wisconsin and NCI groups (PDB entry: 7HVP).[68] The hydroxyl group of the hydroxyethylamine moiety lies between the two catalytic aspartates (D25 and D25′) and makes favorable electrostatic contact in much the same manner as **U-85548e**. The flap water is present and makes hydrogen bonds to the P_1 and P_1' carbonyls and the backbone NH's of Ile-50 and Ile-50′. The P_1 phenyl and P_3 isobutyl groups of the

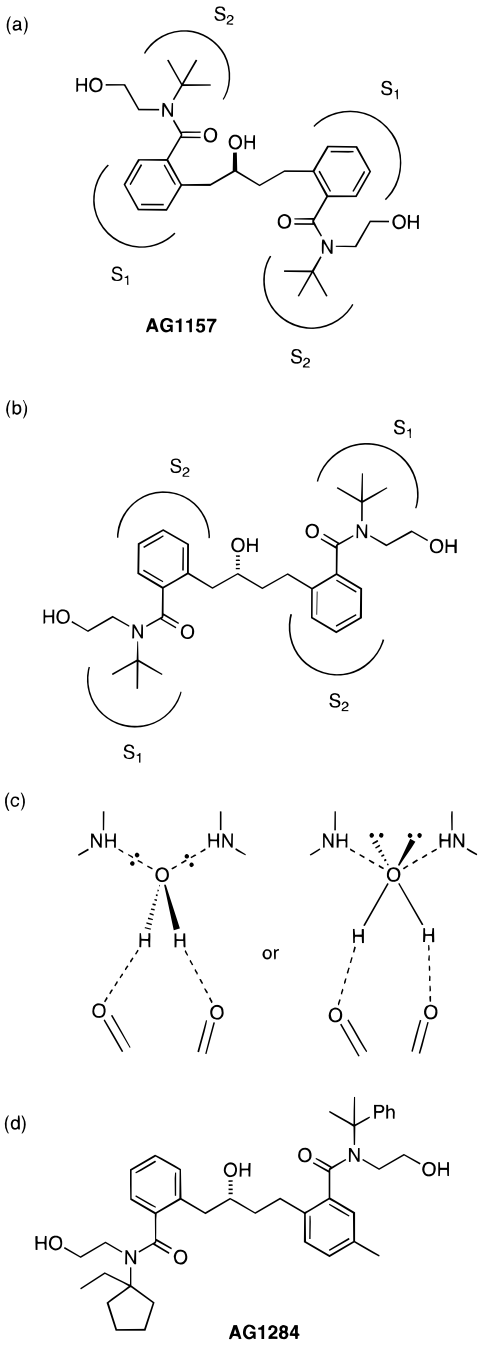


Figure AP*10. Comparison of modeled and observed conformations of the **AG1157** HIVp complex. (a) The hydroxyethyl side chain was designed to incorporate the structural water molecules observed in the **AG1132** complex into the inhibitor. (b) The observed conformation of tertiary amide **AG1157** revealed an inversion in binding mode (see text). (c) In the bound conformation of **AG1157** the flap water interacts with both the protein and the inhibitor with nontetrahedral geometry. (d) Structure is of **AG1284**.

inhibitor make VDW contacts with each other and generate a composite surface. This composite surface makes VDW contacts with the adjacent S_1 and S_3 subsites of the protein. The P_2 asparagine side chain appears to only make VDW contacts with the protein and the P_4 serine side chain makes a hydrogen bond to the side chain of Asp-30. This complex makes essentially the same backbone contacts as does **U-85548e** and will not be discussed further. What

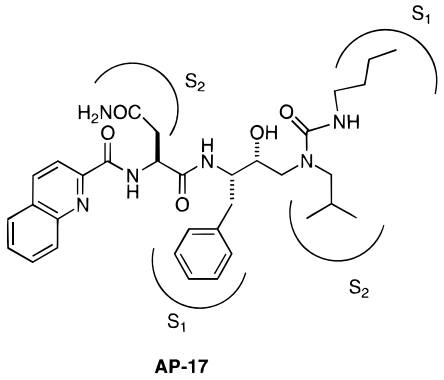


Figure AP*11. An inversion of binding mode was also observed for tertiary urea **AP-17**.

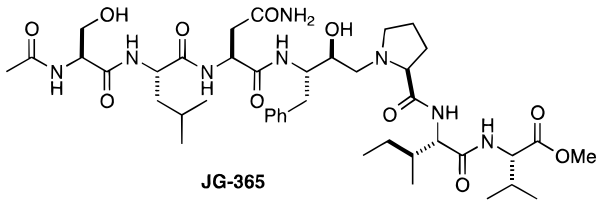


Figure AP*12. Structure of peptidyl hydroxyethylamine inhibitor **JG-365**.

will be focused on is the stereochemical consequences of the central hydroxyl group.

Before 1990, all crystal structures of peptidic inhibitors bound to aspartic proteases had the *S* configuration (or its equivalent) at the central hydroxyl group. This was as expected based on principles derived from studies on pepstatin and pepstatin analogs. Compounds **U-85548e** and **JG-365** were not exceptions to that general rule. For **JG-365** the *S* diastereomer was 85 times more potent an HIVp inhibitor than the *R* diastereomer. In fact, the crystal structure for **JG-365** was obtained from a 50:50 mixture of the *R* and *S* isomers (only the *S* isomer was bound). The backbone chains of **JG-365** and **U-85548e** closely superimpose from P_4 to the methylene unit of central hydroxyethylene or hydroxyethylamine moieties. In addition the P_2' and P_3' groups superimpose. For **JG-365**, the proline ring makes VDW contacts with the S_1 subsite of HIVp. The *S* stereochemistry of the carbon bearing the hydroxyl group allows for simultaneous placement of the hydroxyl group between the catalytic aspartates, optimal interactions for P_1 through P_4 and P_2' through P_3', and the placement of the proline ring in the S_1 and subsite without any strain in the central linking group.

In 1990, a group at Roche UK reported a potent HIVp inhibitor (K_i = 0.12 nM), **Ro 31-8959**, which contained a novel decahydroisoquinoline hydroxyethylamine isostere replacement for the Phe−Pro cleavage site.[69,70] This compound is now the marketed anti-AIDS drug saquinavir. The structure of **Ro 31-8959** was very interesting because it had *R* stereochemistry at the carbon bearing the central hydroxyl group. The *S* hydroxyl epimer of **Ro 31-8959** was 1000 times less potent. This result was quite surprising since, at that time, it was unique among aspartic protease inhibitors and in contrast to the principles derived from studies on pepstatin analogs.

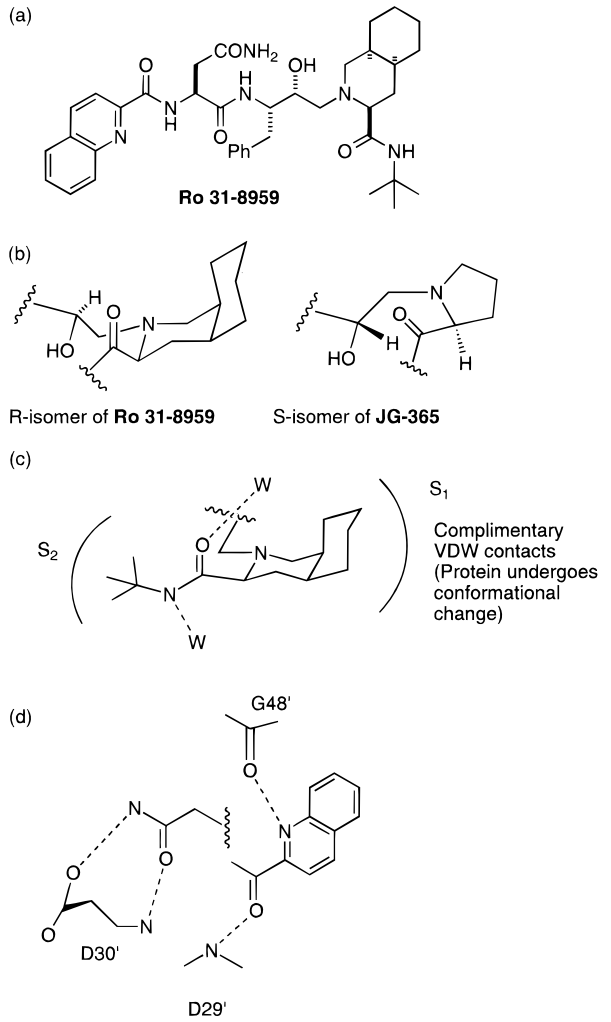


Figure AP*13. (a) Structure of **Ro 31-8959**. (b) The *R* isomer of **Ro 31-8959** and *S* isomer of **JG-365** interact with HIVp by placing the central hydroxyl group between the catalytic aspartates, and place the P_1' group in the adjacent S_1 subsite without inducing strain in the connecting chain. (c) The DIQ P_1' group and *tert*-butyl P_2' group of **Ro 31-8959** make highly complimentary VDW contacts with the protein. The large DIQ group causes a conformation change in the protein. The P_2' amide NH group interacts with the protein through ordered water molecules. (d) The P_2 asparagine group forms hydrogen bonds to Asp-30′ of HIVp and the quinoline amide group forms a hydrogen bond to the backbone NH of Asp-29′. The quinoline ring nitrogen makes a close, presumably unfavorable, contact with the carbonyl of Gly-48′.

The crystal structure of the complex between **Ro 31-8959** and HIVp shows several interesting features (see Figure AP*13):[56,70] (1) The central hydroxyl group is, as expected, located between Asp-25 and Asp-25′. The P_1 and P_1' carbonyl groups make hydrogen bonds to the flap water. (2) The [(4a*S*,8a*S*)-decahydroisoquinin-3(*S*)-carbonyl]oxy (DIQ) group shows excellent steric complimentary to the S_1 subsite and makes extensive VDW contacts. The attached *tert*-butyl amide side chain makes a hydrogen bond from the carbonyl group to the flap water and the *tert*-butyl group makes VDW contacts with the S_2 subsite with its methyl groups completely buried. The NH of the *tert*-butyl amide interacts via highly ordered floor water molecules with the backbone NH of Asp-29 and the backbone carbonyl of Gly-27. (3) The NH_2 group of the P_2 asparagine side chain is located 3.1 Å from the side chain of Asp-30, forming a weak hydrogen bond to the anti-lone pair. (4) The P_1 phenyl group fits into the S_1 subsite and makes hydrophobic contacts. The terminal 2-quinoline group fits into the S_3 subsite and makes a favorable edge to face interaction with the P_1 phenyl group of the inhibitor. The quinoline ring and the adjacent carbonyl are in a low-energy local conformation. It is quite possible that these two groups (phenyl and quinoline rings) interact with each other in solution and induce a bioactive conformation by ligand hydrophobic collapse. (5) The amide NH between the P_2 asparagine and quinoline ring makes a weak 3.3 Å hydrogen bond to the main chain carbonyl of Gly-48 and the ring nitrogen of the quinoline is located 3.1 Å from the main chain carbonyl of Gly-48. The later is presumably an unfavorable electrostatic interaction. The observed conformation of the quinoline is the one in which the electronegative nitrogen atom in the aromatic ring lies anti to the electronegative carbonyl oxygen atom, diminishing an unfavorable electronic interaction. Thus, the free ligand is likely preorganized for binding to the enzyme in its low-energy local conformation. (6) To accommodate the large DIQ ring system the flap residue Pro-81 (and the surrounding residues from Pro-79 to Val-82) undergo a conformational change (compared to the **U-85548e** and **JG-365** complexes) upon binding **Ro 31-8959**. This results in a deviation from C_2 symmetry and generates S_1 and S_1' pockets of different sizes. (7) The favorable protein−ligand interactions discussed above with the central hydroxyl located between the catalytic aspartates are readily accommodated without introducing undue strain in the ligand by having the *R* configuration at the carbon bearing the central hydroxyl group.

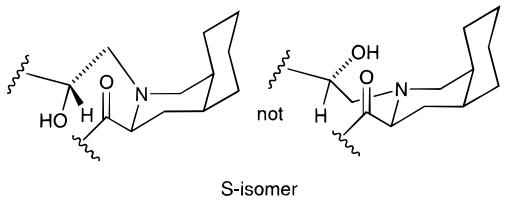
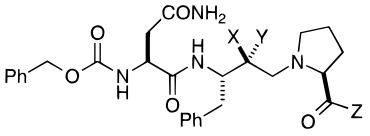
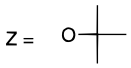


Figure AP*14. The less potent *S* isomer of **Ro 31-8959** binds to HIVp with an inversion of configuration at the tertiary amine.

The crystal structure of the *S* epimer of **Ro 31-8959** has been obtained by the Agouron group.[56] In this structure all the interactions 1−6 above are almost identical to those found in the HIVp complex of **Ro 31-8959**. However, in the *S* diastereomer there has been an inversion of configuration at the tertiary amine. This results in a substituent being axially disposed within the DIQ ring system. (See Figure AP*14.) It appears that the large difference in affinity between the *R* and the *S* epimers is primarily due to an increase in ligand strain for the *S* isomer. A simple modeling exercise can be conducted for the *R* isomer of **Ro 31-8959** where the stereocenter bearing the hydroxyl group is inverted by exchanging the positions of the hydroxyl and hydrogen (generating the *S* isomer). This does not appear to be an unreasonable binding mode. That this binding mode is not observed (hydroxyl not between aspartates) shows that it is higher in energy than the observed

Table AP*2. Structures and IC_{50} Data for Various Hydroxyethylamine Inhibitors



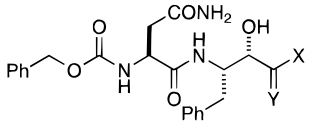
	IC_{50} (nM)
Z = Ile-Val-OMe	
X = OH, Y = H	13.
X = H, Y = OH	>>100.
Z = 	
X = OH, Y = H	300.
X = H, Y = OH	140.

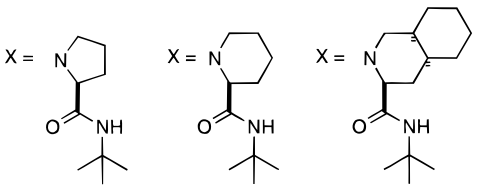
binding mode (hydroxyl between aspartates with resulting ligand strain). The observed structure of the *S* isomer and the difference in activity between the *R* and *S* isomers provide experimental support that placing the hydroxyl group between the catalytic aspartates is worth more than 4 kcal mol^{-1}.

Other data also supports the notion that the optimal hydroxyl group stereochemistry will be the one that allows simultaneous placement of the central hydroxyl group between the catalytic aspartates and optimal interactions of the P_1, P_2, P_1', and P_2' groups with the protein without inducing undue strain in the ligand. The Roche[70] and Wisconsin[71] groups have generated data in a series of hydroxyethyl amine inhibitors derived from Phe-Pro (Table AP*2). In these examples, inhibitors related to **JG-365**, with peptidyl P_2' and P_3' groups, the *S* isomer is the more potent inhibitor. However, for inhibitors related to **Ro 31-8959** terminating with a *tert*-butyl P_2' group, the two isomers have nearly equivalent activity, with the *R* isomer being slightly more potent. In the case of **Ro 31-8959**, with a DIQ P_1' group, the *R* isomer becomes preferred by a factor of 1000.

The Syntex group[72] examined α-hydroxy amide and α-keto amide inhibitors based upon **JG-365** and **Ro 31-8959**. A different SAR trend was observed for this series of compounds. (See Table AP*3.) In the case of close analogs of **Ro 31-8959**, increased potency was observed as the P_1' group was increased in size from proline to piperidine to DIQ. For the α-hydroxy amides, activity decreased as the P_1' group was increased in size. It has been proposed that ligand conformational effects may be responsible for this loss in potency. As observed for the *S* isomer of **Ro 31-8959**, the DIQ ring system binds to the protein in a chair−chair *cis*-decalin conformation with the *tert*-butyl amide substituent equatorial. For α-hydroxy amides with an equatorial *tert*-butyl amide group, a repulsive interaction between the *tert*-butyl amide group and the adjacent *N*-acyl group occurs as a result of $A_{1,3}$ strain. (See Figure AP*15.) Therefore, either a high-energy conformation binds to the protein or a relaxed conformation binds to the protein making less than optimal interactions. No crystal structures have been reported for these or related inhibitors. For the proline-based inhibitor, the hydroxyl stereochemistry for the more potent isomer was the same (note: the priority rules reverse the *R* and *S* designations) as in the hydroxyethyl amine series. Interestingly, the α-keto amide retains significant activity compared to the most potent hydroxy epimer in the α-hydroxy amide series.

Table AP*3. Structures and IC_{50} Data for α-Hydroxy Amide Inhibitors





Y = H,H	210 nM	18 nM	< 2.7 nM
Y = O	7 nM	26 nM	80 nM
R-isomer*	3100 nM		
Ketoamide	20 nM		

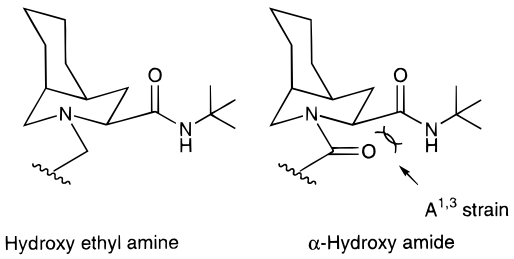


Figure AP*15. While no structures of the hydroxy amide compounds in Table AP*3 have been reported, unfavorable $A_{1,3}$ strain would result if they bound similar to **Ro 31-8959**.

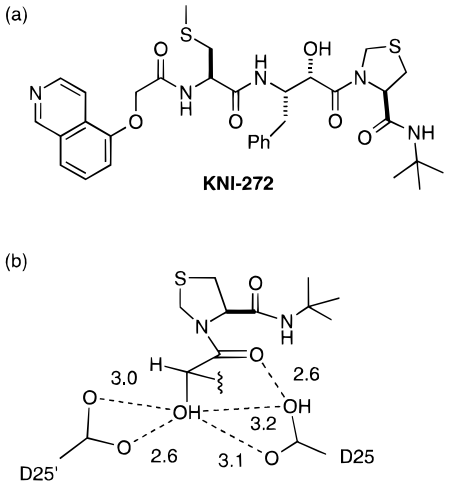


Figure AP*16. (a) Structure of **KNI-272** and (b) interactions of α-hydroxyl amide group with catalytic aspartates.

The structure of a related inhibitor, **KNI-272** (K_i ≈ 5 pM), bound to HIVp has been reported (PDB entry: 1HPX; see Figure AP*16).[73] This structure shows that the central hydroxyl group interacts more strongly with Asp-25′ than it does with Asp-25; however, the amide carbonyl oxygen forms a short

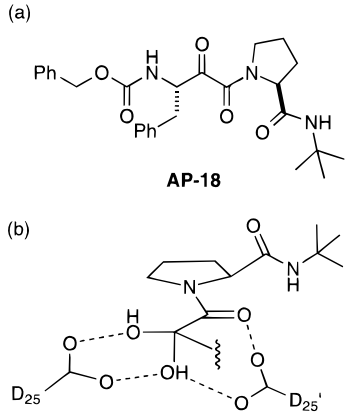


Figure AP*17. (a) Structure of **AP-18** and (b) interactions of α-keto amide group with catalytic aspartates.

hydrogen bond (2.6 Å) with one of the oxygens of Asp-25. NMR experiments have been reported which assign the protonation state of each of the catalytic aspartates. Asp-25, which forms a hydrogen bond to a carbonyl, is protonated and Asp-25′ which strongly interacts with the inhibitor hydroxyl group is charged and unprotonated. In this case, there is experimental evidence that the inhibitor binds to the −1,0 protonation state of the enzyme. In this NMR experiment, the enzyme was asymmetric on the NMR time scale since resonances for each protein monomer were observed.[74]

A group at Scripps[75] has also described pyrrolidine-containing α-keto amide HIVp inhibitors. In their studies they report inhibitor **AP-18** (K_i = 214 nM) and a crystal structure of its complex with HIVp. The structure shows that the **AP-18** is bound to HIVp in its hydrated state. (See Figure AP*17.) The *pro-S*-hydroxyl is located between the catalytic aspartates and makes hydrogen bonds to each of them. The *pro-R*-hydroxyl interacts with only one of the two aspartates (D25) and the amide carbonyl makes a hydrogen bond to D25′. The flap water is present; however, its location is rather asymmetric. The related α-hydroxy amide inhibitors were prepared. The *S* isomer had a K_i = 2 μM and the *R* isomer was less potent with a K_i = 300 μM.

This structure raises an interesting question. Is the inhibitory species the hydrated or the unhydrated keto amide? If it is the hydrated species then the inhibitor displaces the catalytic water molecule. If the amide is the inhibitory species then it must be hydrated in the active site by the catalytic water of the enzyme. For compound **AP-18**, a ^{13}C NMR experiment in DMSO-d_6/D_2O (5:1) determined that, in the presence of water, the ketone group of **AP-18** remains unhydrated even after incubation for 24 h. In addition, inhibitor **AP-18** does not exhibit time dependent inhibition. These results suggest that the unhydrated ketone is the inhibitory species, and that hydration takes place within the enzyme active site. The resulting hydrated amide is then stabilized by hydrogen bonds to the catalytic aspartates. If hydration takes place within the enzyme, then the catalytic water is not displaced into bulk solvent but rather incorporated into the inhibitor. In this case the entropic gain of the water displacement process is not to be expected.

The structure of **Ro 31-8959**, a (*R*)-hydroxyethylamine isostere, and the structure of its complex with HIVp provided a new starting point for further inhibitor design.

A collaborative effort between Lilly and Agouron used an iterative structure-based design cycle to prepare new inhibitors of HIVp based upon the Phe−Pro cleavage site. These studies ultimately resulted in **AG1343** that is presently in clinical trials.[76−80] During the first part of this cycle, an ortho-substituted benzamide was designed as a DIQ (proline) mimic. This resulted in **LY289612** (K_i = 1.5 nM). The structure of the complex between **LY289612** and HIVp is similar in many respects to the **Ro 31-8959** HIVp complex. The significant differences include a 180° rotation about the C=O quinoline bond (possibly due to crystal packing) and substitution of the DIQ ring with a benzene ring. The later substitution results in an HIVp conformation for residues 79−82 similar to that observed in **U-85548e** and **JG-365**. The ortho substituent adjacent to the benzamide favors the amide to be out of conjugation and allows hydrogen-bond formation with the flap water.

The next phase in the design process required the substitution of the P_2 and P_3 groups with nonpeptidyl groups. It was noted that the P_1 and P_3 groups of the **LY289612** made VDW contacts with each other and suggested that a larger P_1 group could span both the S_1 and S_3 subsites. This resulted in an *S*-naphthyl P_1 substituent. Inhibitor **AP-19** is a potent inhibitor with IC_{50} = 1.1 nM. The crystal structure of **AP-19** shows that the large *S*-naphthyl group occupies the adjacent S_1 and S_3 subsites. To accommodate the large substituent at P_1/P_3, the protein undergoes a conformational change in flap residue Pro-81 (and the surrounding residues from Pro-79 to Val-82). This results in a deviation from C_2 symmetry and generates S_1 and S_1' pockets of different sizes. This is similar to the protein changes observed in the **Ro 31-8959** complex. Inhibitor **AP-20** is about 1 order of magnitude more potent than **Ro 31-8959**, and is one of the most potent HIVp inhibitors reported.[78] After several iterative design cycles, the 2-methyl-3-hydroxybenzamide P_2 group was invented resulting in **AG1254** (K_i = 2.7 nM). The structure of the complex between **AG1254** and HIVp (see Figure AP*18) shows, as expected, the hydroxyl group between the two aspartates and hydrogen bonds to the flap water. The *S*-naphthyl group at P_1 spans the S_1/S_3 subsites and the protein undergoes a conformational change as observed for **AP-19**. The phenol of the P_2 group forms a hydrogen bond (2.9 Å) with an anti-lone pair of Asp-30 and a longer (3.4 Å) electrostatic interaction with the backbone NH of Asp-30. The ortho methyl group is buried in the S_2 subsite and favors the amide to be out of conjugation and allows hydrogen bond formation with the flap water. Structural modifications of **AG1254** to improve its pharmacological profile resulted in **AG1343**. The complex structure of **AG1343** is as expected. The larger DIQ group results in a conformational change in flap. The *S*-phenyl group in **AG1343** makes extensive hydrophobic contacts with an S_1-S_3 subsite and causes a conformational change in the protein. The S_1 and S_2 aromatic rings make an edge-to-face

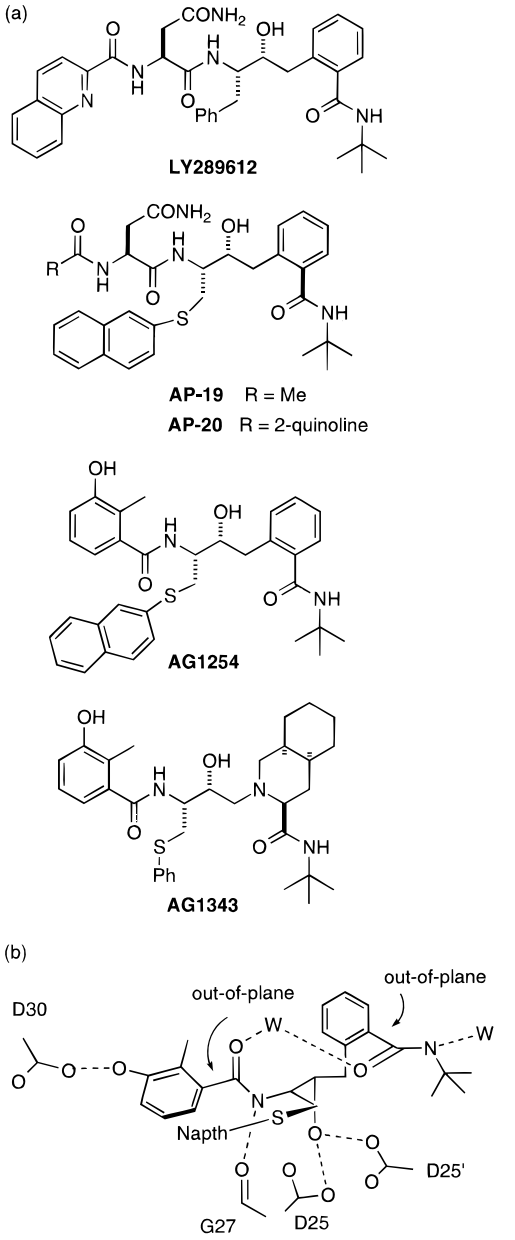


Figure AP*18. (a) Structures of various inhibitors. Compound **AG1343** is the marketed drug nelfinavir. (b) Key interactions of **AG1254** with HIVp. The ortho substituents on both benzamides allow the carbonyl group to be out-of-plane with the aromatic ring. This allows both carbonyl groups to interact with the flap water. The phenol of the P_2 group forms a hydrogen bond with Asp-30.

interaction in the complex. It is possible that these two groups also interact with each other in solution and induce a bioactive conformation by ligand hydrophobic collapse. A derivative of **AG1343** not having the sulfur atom (i.e., *S*-phenyl = phenyl) is a weaker enzyme inhibitor by an order of magnitude. Thus, the sulfur atom plays an important role. In this structure both S_1 subsites are enlarged relative to **U-85548e** and **JG-365**; these two subsites are still different in size with the one adjacent to the *S*-phenyl group being larger. **AG1343** is now the marketed anti AIDS drug nelfinavir.

The Vertex group reported the results of a focused structure-based design program that sought to maintain subnanomolar inhibitor potency while reducing inhibitor size. The novel inhibitor, **VX-478**, $K_i = 0.6$ nM, MW = 506, emerged from these studies.[81] This compound has two novel structural features, a sulfonamide group in the P_2' position and a tetrahydrofuran in the P_2 position. Compound **VX-478** is presently undergoing clinical trials.

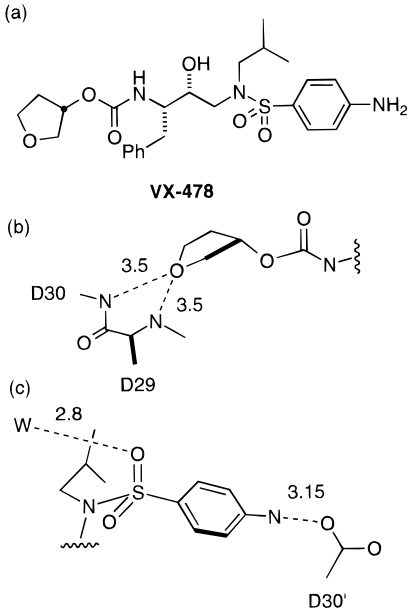


Figure AP*19. (a) Structure of **VX-478**. (b) The oxygen of the tetrahydrofuran ring makes weakly polar interactions with the backbone NH's of both Asp-29 and Asp-30. (c) The sulfonamide is in a low-energy conformation and interacts with the flap water. The *p*-NH_2 substituent interacts with Asp-30′.

Compound **VX-478** has a central (*R*)-hydroxyethylamine isostere. The structure between **VX-478** and HIVp (PDB entry: 1HPV; see Figure AP*19) shows the hydroxyl group located between the two aspartates and hydrogen bonding to the flap water from the P_1 carbonyl group as well as from one of the sulfonamide oxygens. The other sulfonamide oxygen is partially buried in a hydrophobic environment and 3.4 Å from the flap water. Compound **VX-478** does not form a hydrogen bond to either Gly-27 or Gly-27′; in both cases the carbonyl is partially buried by the inhibitor and partially solvent exposed. The conformation observed for the benzenesulfonamide appears to be close to a local minimum on the basis of *ab initio* calculations and examination of small molecule crystal structures. The P_2' aromatic makes hydrophobic contacts with an S_2 subsite, and the amino substituent makes a hydrogen bond to Asp-30′. The backbone chain near the central hydroxyl has a conformation similar to that in **Ro 31-8959**. In this complex, it appears that extensive hydrophobic interactions play a major role in stabilizing the complex. The P_2 tetrahydrofuran ring is in an energetically preferred pseudoaxial conformation, and the ring oxygen makes weak electrostatic interactions with the backbone NH's of Asp-29 and Asp-30 (3.4 and 3.5 Å, respectively).

The Merck group has also reported on 3-tetrahydrofuran urethanes as high affinity P_2 ligands (see Figure AP*20).[82,83] Their studies in hydroxyethylene and hydroxyethylamine series have shown that the *S* configuration provides a 4-fold increase in potency relative to the *R* configuration and that substitution

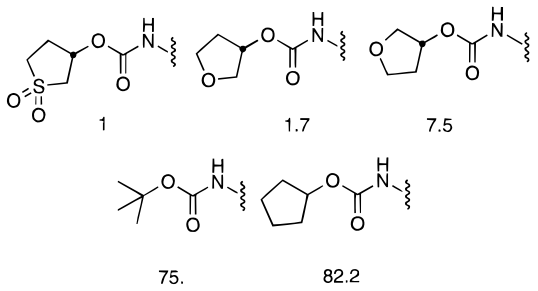
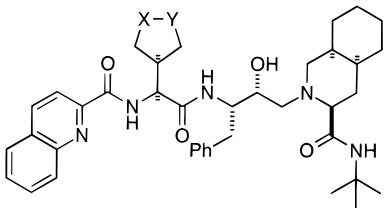


Figure AP*20. Relative IC_{50} values for various P_2 carbamate groups.



AP-21 X = CH_2, Y = O
AP-22 Y = CH_2, X = O

Figure AP*21. Structures of various inhibitors.

of the ring oxygen with a methylene results in a 50-fold loss in potency. The Merck group has also reported on the extension of this work to cyclic sulfolanes. These P_2 ligands show a 1.7-fold improvement in affinity relative to the tetrahydrofurans.[84] While a crystal structure between a simple cyclic sulfolane and HIVp has not been reported, it is tempting to speculate that the weak electrostatic interactions with the backbone NH's of Asp-29 and Asp-30 made by the ring oxygens of **VX-478** are better optimized with the cyclic sulfolane. These SAR studies have shown that careful optimization of protein−ligand electrostatics in a P_2 group can contribute about 2.5 kcal mol^{-1}.

The Merck group has also shown that 3′-tetrahydrofuranylglycine is a novel, unnatural amino acid surrogate for asparagine in HIVp inhibitors. Compound **AP-21**, an analog of **Ro 31-8959** is a very potent inhibitor (IC_{50} = 54 pM) and was reproducibly 4-fold more potent than **Ro 31-8959** as a competitive HIVp inhibitor. The epimeric **AP-22** was 100-fold less effective than **AP-21**.[85] (See Figure AP*21.)

3. C_2-Symmetric, Pseudosymmetric, and Hydroxylamine Pentanamide-Based Inhibitors

As soon as the structure of HIVp was determined, it was recognized that C_2 symmetrical molecules had potential as inhibitors of this enzyme.[86] The key to this design strategy is that the symmetry of the inhibitor and the enzyme coalign. In 1990, the Abbott group was the first to describe pseudosymmetric and C_2-symmetric inhibitors of HIVp. The inhibitors they described were symmetric with respect to the "unprimed" side of a substrate. In their initial report they described the design, activity, and crystal structure of the pseudosymmetric inhibitor **A-74704** (K_i = 4.5 nM; PDB entry: 9HVP; see Figure AP*22).[87]

The structure of this complex showed the expected features of the central hydroxyl group located between Asp-25 and Asp-25′ and hydrogen bonds

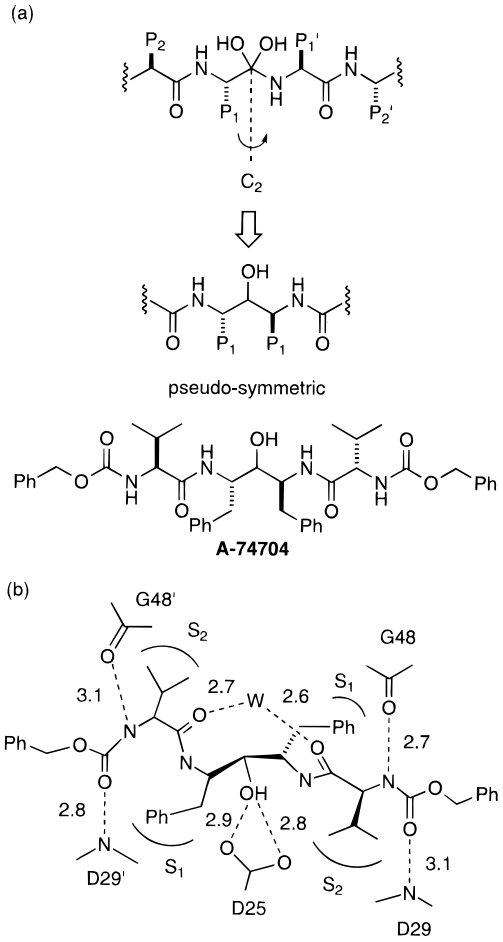


Figure AP*22. (a) A C_2-symmetry operation through the carbon atom of a tetrahedral intermediate suggests pseudosymmetric inhibitors such as **A-74704**. (b) Schematic of interactions of **A-74704** with HIVp.

between the central carbonyls of the inhibitor and the flap water. The central hydroxyl lies within 0.2 Å of the molecular 2-fold axis and the flap water lies within 0.2 Å of that axis. The enzyme maintained its C_2 symmetry and, with the exception of the terminal Cbz groups, so did the inhibitor. The symmetry axes of both protein and inhibitor coaligned. Hydrogen bonds between the protein and the inhibitor were similar to those seen in asymmetrical peptide-like compounds. Hydrogen bonds from the backbone NH of Asp-29 to the inhibitor urethane carbonyl and hydrogen bonds between the backbone carbonyl of Gly-48 and the urethane NH were observed. These hydrogen bonds occur twice in the symmetrical inhibitor and are related to each other by a C_2 axis. Hydrogen bond formation from the protein to the flap water requires a major element of protein asymmetry; however, the 2.8 Å resolution of the structure made interpretation of that region ambiguous.

The Abbott group also reported on C_2 symmetric diols where the symmetry axis lies on the bond between the two alcohols.[88] All three possible diol stereoisomers (the C_2 symmetric *R*,*R* and *S*,*S* and the nearly symmetric *S*,*R*) were individually prepared and evaluated. High-resolution crystal structures were obtained for the complexes of each of these isomers (C_2-symmetric *S*,*S* diol: **A-76928**, K_i = 77 pM [PDB entry: 1HVK], C_2-symmetric *R*,*R* diol:

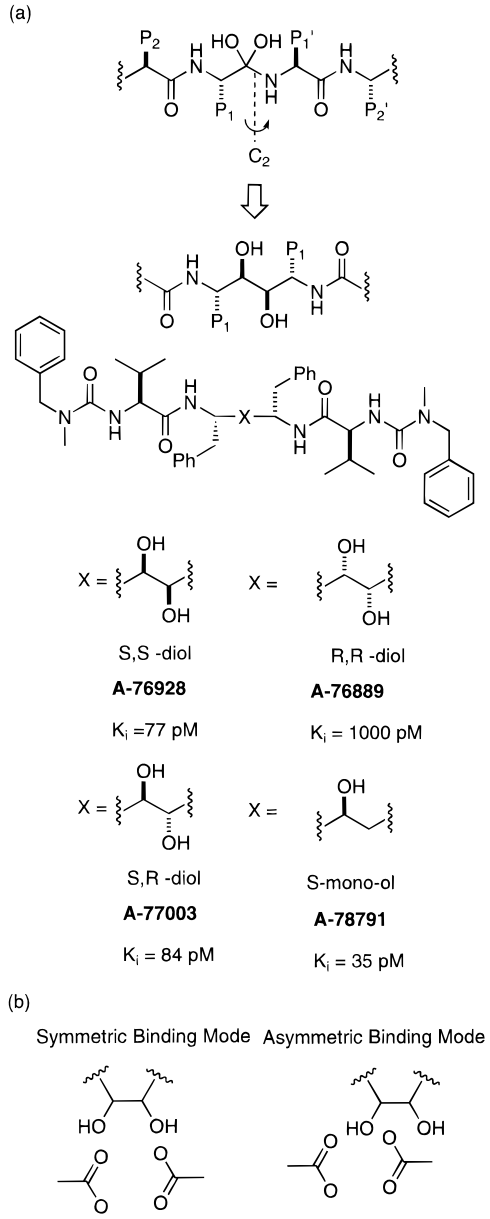


Figure AP*23. (a) A C_2-symmetry operation through the midpoint of the C−N peptide bond suggests C_2-symmetric inhibitors such as **A-76928** and **A-76889**, and the pseudo-symmetric inhibitors **A-77003** and **A-78791.** (b) Do C_2-symmetric compounds binds to the aspartates symmetrically, in which each hydroxyl interacts equally with each aspartate, or asymmetrically where only one hydroxyl interacts strongly with both aspartates?

A-76889, K_i = 1000 pM [PDB entry: 1HVL], the nearly symmetric *S,R* diol: **A-77003**, K_i = 84 pM [PDB entry: 1HVI) as well as a structure of related deshydroxy analog (**A-78791**, K_i = 35 pM [PDB entry: 1HVJ].[89] For all four inhibitors, the P_1, P_2, and P_3 groups bound almost identically and exhibited features common to most peptidomimetic HIVp inhibitors, and all made hydrogen bonds to the flap water. The most important issue to be addressed was whether the diols would bind asymmetrically, where only one hydroxyl was between the aspartates, or symmetrically, where both hydroxyls would interact equally with the aspartates. (See Figure AP*23.)

The more potent C_2-symmetric *S,S* diol, **A-76928**, binds to HIVp in a novel symmetrical mode, in which each hydroxyl group makes nearly equivalent interactions with a single aspartate. (See Figure AP*24.) Each hydroxyl oxygen is 2.7 Å from the nearest oxygen on the adjacent carboxyl group and can make three additional hydrogen bonds within 3.1−3.3 Å with the remaining oxygens for a total of eight potential OH−O contacts. The C_2 axis of the enzyme passes within 0.1 Å of the midpoint of the diol bond. Compound **A-76928** is truly a symmetrical inhibitor of HIVp.

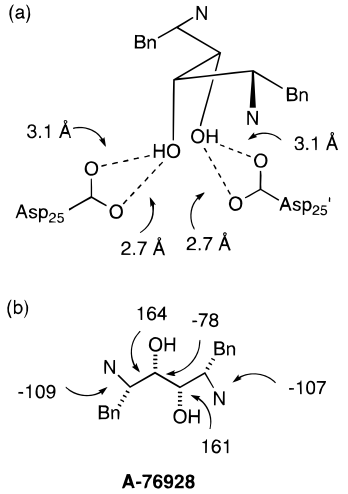


Figure AP*24. (a) Interactions of *S,S* diol unit of **A-76928** with catalytic aspartates and (b) backbone torsion angles of bound conformation of **A-76928**.

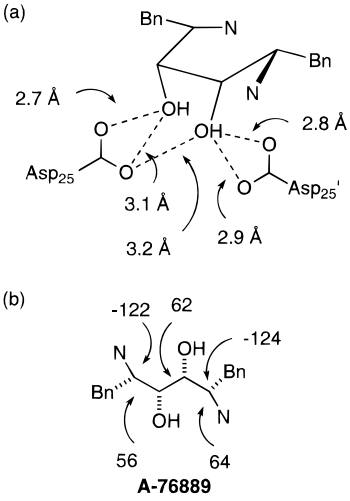


Figure AP*25. (a) Interactions of *R,R* diol unit of **A-76889** with catalytic aspartates. (b) Backbone torsion angles of bound conformation of **A-76889**, the poorer inhibition of the *R,R* isomer is likely due to less optimal interactions of the diol with the catalytic aspartates and, possibly, ligand strain of the inhibitor backbone.

The less potent C_2-symmetric *R,R* diol, **A-76889**, binds to HIVp in a novel asymmetrical mode. The two hydroxyls both interact with the catalytic aspartates, one more intimately than the other. The more central hydroxyl interacts with both aspartates, while the more distal one can only interact with one of the aspartates. The most noticeable difference between it and **A-76928** is the conformation of the central unit. (See Figure AP*25 and list of torsion angles.) Comparison of the pseudosymmetric **A-74704**, and the asymmetric **A-77003** and **A-78791** show that the central hydroxyl oxygens of all three inhibitors are located at the same position in the active site. The

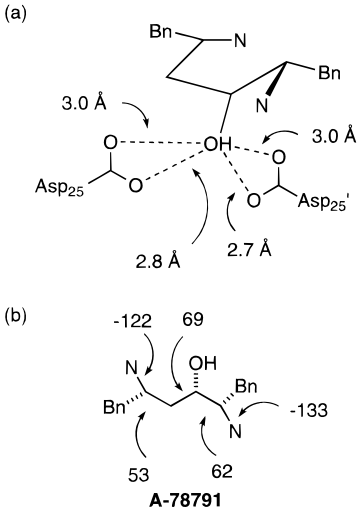


Figure AP*26. (a) Interactions of alcohol unit of pseudosymmetric **A-74704** with catalytic aspartates and (b) the backbone torsion angles of bound conformation of **A-74704** are very similar to the *S*,*S* diol inhibitor **A-76928**.

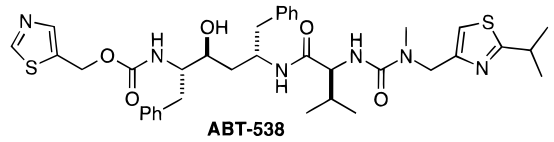


Figure AP*27. Structure of **ABT-538**, the marketed drug ritonavir.

C_2-symmetric **A-76889** places its more central hydroxyl in a different location. The weaker potency of **A-76889** compared to **A-77003** and **A-78791** is likely due to less than optimal interactions with the two aspartates.

For the deshydroxy compound **A-77003**, as expected, an asymmetric binding mode was observed. (See Figure AP*26.) The conformation of the central unit of this compound is very similar to that of the more potent *S*,*S* diol **A-76928**. For the asymmetric *S*,*R* diol, **A-77003**, an asymmetric binding mode was also observed very similar to the deshydroxy compound. For these two asymmetric compounds and the pseudosymmetric **A-74704**, the orientation of the central hydroxyl group between the aspartates is similar to the orientation seen in many asymmetric compounds.

The Abbott group designed novel HIVp inhibitors on first principles. Final optimization of these compounds with respect to pharmacological properties led to the asymmetric **ABT-538** that is presently the marketed AIDS drug ritonavir.[90] (See Figure AP*27.)

A phosphinate-based C_2-symmetric inhibitor, **SB204144**, and its complex with HIVp were reported by the SmithKline group.[91] The K_i reported for this compound was 2.8 nM and when compared directly with **A-74704** was found to be 82 times more potent. Compound **SB204144** is weakly antiviral, presumably due to poor cellular penetration. This brings back the important theme that a useful drug candidate, in addition to being a potent enzyme inhibitor, must have desirable pharmacological properties.

The overall structure of this complex is very similar to the **A-74704** complex. (See Figure AP*28.) In this complex, the flap water and the O−P−O atoms lie in a plane and this plane lies on the symmetry axis.

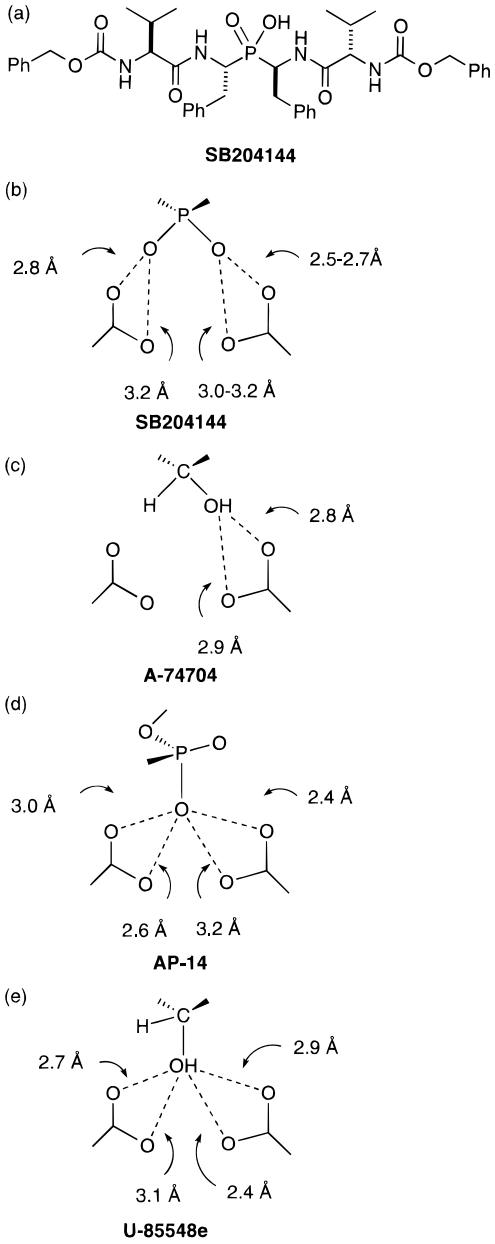


Figure AP*28. (a) Structure of **SB214144** (comparison of interactions of different inhibitors with catalytic aspartates), (b) C_2-symmetric **SB214144** with HIVp, (c) pseudosymmetric **A-74704** with HIVp, (d) asymmetric **AP-14** with penicillopepsin, and (e) asymmetric **U-85548e** with HIVp.

Each oxygen of the phosphinate interacts with a single aspartate, making one short (2.5−2.8 Å) and one long (3.0−3.2 Å) oxygen−oxygen contact. The close contacts between the phosphinate and carboxylates would seem to be most consistent with the three functional groups sharing a −1 charge. One limitation of protein crystallography is the inability to determine the positions of hydrogen atoms.

Phosphinate **SB204144**, that has an experimentally determined pK_a of 3.1, is a potent (K_i = 2.8 nM) inhibitor of HIVp at pH 6.0. This suggests that either the ionized form is a potent inhibitor, binding to the 0, 0 form of the enzyme, or that the neutral form of the phosphinate must be a picomolar inhibitor binding to the −1,0 state. An analysis of pH dependence on K_i failed to resolve this question.

The protonation state of aspartic proteases is usually represented as having one aspartate deprotonated and the other protonated (−1,0 state). A study of pH rate studies of model substrates for HIVp has show that the two aspartate groups have distinct pK_a values of 3.1 and 5.2. In that study it was demonstrated that a neutral statine-based inhibitor bound only to the monoprotonated −1,0 state while a charged "reduced amide" inhibitor bound only to the doubly deprotonated (−1,−1) state.[27] A discussion based on molecular dynamics simulations speculates that the active-site protonation state will be dependent on the local environment near the aspartates and will be different for different inhibitors.[92]

Compound **SB204144** was determined to be 80 times more potent than the secondary alcohol **A-74704**. This shows that the observed close contacts made by the phosphinate are extremely favorable. A charged phosphinate presumably pays a much greater desolvation penalty than a secondary alcohol; therefore, the protein−ligand interaction between the phosphinate and the catalytic aspartates must be very strong.

It is instructive to compare the location of the oxygens between the catalytic aspartates in the HIVp complexes of **SB204144**, **A-74704**, and **U-85548e** and the penicillopepsin phosphonate inhibitor **AP-14**. (See Figure AP*28.) In **A-74704**, the location of the hydroxyl group is coincident with the location of one of the phosphinate oxygens of **SB204144**. In **U-85548e** the location of the hydroxyl group lies within the O−P−O plane of **SB204144** and is approximately midway between the two phosphinate oxygens. A slight conformational difference in one of the aspartates is observed in the **U-85548e** complex. The structure of a phosphonate transition-state analog (**AP-14**) complexed to penicillopepsin shows that the two phosphonate oxygens interact with the catalytic aspartates unequally. The SmithKline group discussed mechanistic implications for the symmetrical binding of the phosphonate **SB204144**.[91]

An alternative class of pseudosymmetrical HIVp inhibitors has been independently reported by groups at both Merck[93] and Lederle.[94−96] Unlike the Abbott class of compounds, this class of inhibitors is symmetric with respect to the "prime side" of a substrate. (See Figure AP*29.) Both groups have described a crystal structure for a member of this series that contain non-amino acid-derived P_2' groups.

For the complex between both **L-700,417** (PDB entry: 4PHV) and **AP-23** (R = 2-quinolyl) with HIVp the central hydroxyl group and the two carbonyl groups of the central unit displace the catalytic water and hydrogen bond to the flap water, respectively. The P_1' phenyl groups each make hydrophobic contacts with the enzyme S_1 subsites, and the indan phenyl of **L-700,417** and the cyclohexyl of **AP-23** make hydrophobic contacts with the enzyme S_2 subsites. For **AP-23**, the carbonyls of the amide groups linking the two cyclohexyl rings to the quinoline groups form hydrogen bonds to the backbone NH's of Asp-29 and Asp-29′. In addition, these carbonyls are close (3.0 Å) to the carboxylate of Asp-29 and Asp-29′. Compared to **Ro 31-8959**, this carbonyl is 1.2 Å closer to the side chain carboxylate

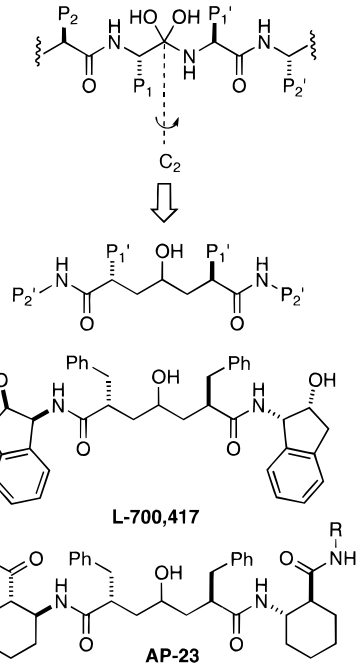


Figure AP*29. A C_2 symmetry operation through the carbon atom of a tetrahedral intermediate also suggests pseudosymmetric inhibitors such as **L-700,417** and **AP-23** which are symmetric with respect to the "prime side" of a substrate.

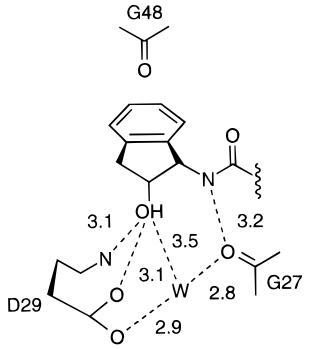


Figure AP*30. Interaction of indanol P_2' group with HIVp. In addition to the hydrogen-bond interactions shown, the phenyl group makes extensive VDW contacts with the protein.

of Asp-29. The amide NH's make less than optimal interactions with the backbone carbonyl of Gly-48 and Gly-48′ (3.4 and 3.6 Å). These linking amide groups presumably do not make optimal electrostatic interactions with the protein.

The indanol ring system present in **L-700,417** is a novel P_2' group that simultaneously provides favorable hydrophobic and electrostatic interactions with HIVp. (See Figure AP*30.) The indanol hydroxyl forms hydrogen bonds with the protein and hydrogen bonds with structural water molecules. The hydroxyl groups of the indanols also are between 3.0 and 3.1 Å of Asp-29 (and Asp-29′). This interaction is presumably quite favorable. Two symmetry-related water molecules donate hydrogen bonds to the carbonyl oxygen of G27 (and G27′) and to the hydroxyl group of the indanol. The carbonyl oxygen of G48 on each HIVp subunit is buried upon complex formation and is not stabilized by hydrogen bonding to either the inhibitor or solvent, although the dipole is aligned directly above the dipole of the indanol

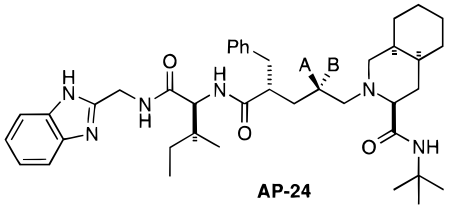


Figure AP*31. Both diastereomeric hydroxylamine pentanamide inhibitors **AP-24** (A = OH, B = H and A = H, B = OH) are potent HIVp inhibitors.

hydroxyl group. This occurs by a slight conformational adjustment of the flap.

Comparison of **L-700,417** and **AP-23** shows some similarities in the two structures as well as some differences. While both compounds have identical heptanediamide central units, they differ slightly in the bound conformation of this group. The phenyl P_1 groups, the central hydroxyl, and central carbonyl oxygens all overlap with each other despite this difference. The major consequence of this different conformation is that the cyclohexyl groups in **AP-23** and the phenyl groups of the indanes of **L-700,417** overlap. That these two groups overlap strongly suggests that burying complementary hydrophobic surfaces is a major driving force in complexation. In addition, the terminal carbonyl oxygens of **AP-23** and the hydroxyl groups of indane **L-700,417** also overlap. Compounds related to **AP-23** require large hydrophobic P_3 groups to be potent enzyme inhibitors while **L-700,417**, which completely lacks a P_3 group, is a very potent enzyme inhibitor. **L-700,417** is 4 orders of magnitude more potent than **AP-23** (R = CH_3). In this example, less than optimal protein−ligand electrostatics made by **AP-23** must be compensated for by additional protein−ligand VDW contacts.

Model building experiments and recognition of the symmetric nature of HIVp led both the Lederle[97,98] and the Merck[99,100] groups to independently design a novel class of hydroxylamine pentanamide inhibitors. The Lederle group prepared **AP-24** (Figure AP*31) as a separable mixture of diastereomers at the carbon bearing the hydroxyl group. It was observed that both isomers were potent inhibitors (*S* isomer [A = H, B = OH] K_i = 7.1 nM; *R* isomer [A = OH, B = H] K_i = 2.7 nM; note *R* and *S* designations are different from the hydroxyethylamine isosteres due to the priority rules) of HIVp. This result was unexpected based on the structure of **Ro 31-8959**. A crystal structure of the *S* isomer was obtained and was as expected for a bidirectional hybrid compound containing the DIQ group at one end and a dipeptide mimic at the other end.[101] Crystals suitable for structure determination have not been obtained for the more interesting *R* isomer. On the basis of the different conformations observed for the linking heptanediamide chains of **L-700,417** and **AP-23**, it is tempting to speculate that the high potency observed for the *R* isomer is due to a relatively flexible linking group that simultaneously allows for optimum interactions on both sides of the inhibitor and placing the central hydroxyl between the two aspartates without introducing strain in the linking unit.

The Merck group reported a stereoselective synthesis of **AP-25**, and it was reported to be a potent

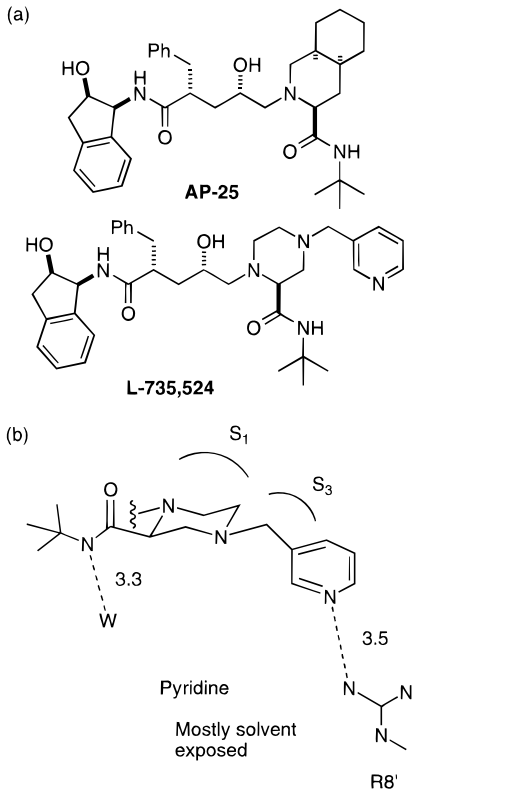


Figure AP*32. (a) Structures of hydroxylamine pentanamide inhibitors. Compound **L-735,524** is the marketed AIDS drug indinavir. (b) Interactions of P_1' group of **L-735,-524** with S_1 and S_3 subsites. The pyridine ring is significantly solvent exposed.

HIVp inhibitor (K_i = 7.8 nM). They have also reported that one member of this class, **L-735,524**, is a potent (K_i = 0.56 nM) and orally bioavailable anti HIV compound. This compound is the marketed AIDS drug indinavir.

A cocrystal structure of **L-735,524** with two strains of HIV-1 protease, NY5 and ROD isolates, and HIV-2 protease, ROD isolate, (PDB entries: 1HSG, 1HSH, 1HSI; see Figure AP*32) have been described.[102] These compounds bound as expected with the CH_2-Py groups in the S_3 subsite. This group makes some VDW contacts with the protein but is somewhat solvent exposed. Modification of the P_3' groups did not generally affect inhibitor potency; however, modifications at this site allowed for the adjustment of pharmacokinetic properties. The previously discussed indanol ring system is present as a P_2' group. Unlike **Ro 31-8959** and **AG1343**, the substituted piperazine P_1' group does not cause a conformational change in the protein in the vicinity of Pro-81.

During random screening, the Glaxo group discovered that a C_2-symmetric penicillin derived compound **AP-26a** was a potent HIVp inhibitor (IC_{50} = 3.0 nM). (See Figure AP*33.) The crystal structure of this complex has been described.[103,104] The two central carbonyls make interactions with the flap water, the two thiazolidine rings each occupy an S_1 subsites and the benzyl groups each occupy an S_2 subsite. There is no water located between the active site aspartates. The aspartates interact weakly with the HN's of the central ethylene diamide. Increasing the size of the central linker from two methylenes to three resulted in more than 2 orders of magnitude

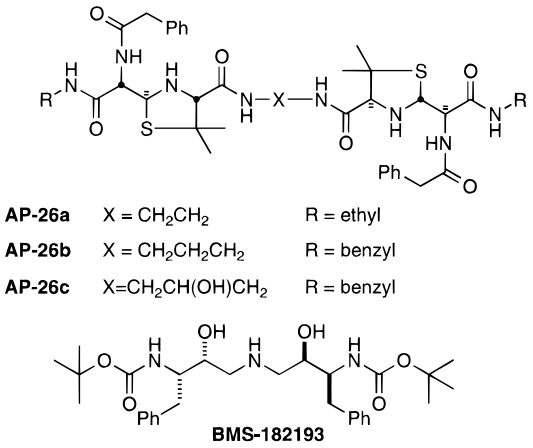


Figure AP*33. Structures of C_2-symmetric HIVp inhibitors.

loss in potency. Introducing a hydroxyl group in the three carbon methylene chain resulted in an increase in potency of 110-fold (2.8 kcal mol^{-1}).

C_2-Symmetrical aminodiols have been described by a group at Bristol-Myers Squibb.[105,106] **BMS-182193** has a K_i = 125 nM. A crystal structure between **BMS-182193** and HIVp shows that the flap water bound symmetrically to each Boc carbonyl, and that each hydroxyl group forms a hydrogen bond symmetrically with each catalytic aspartate.

4. Inhibitors That Replace the Flap Water

A common feature of all the complexes described so far is the presence of the flap water hydrogen bonding to both the protein and the inhibitor. The DuPont Merck group[107] hypothesized that incorporation of binding features of the flap water into an inhibitor would be energetically favorable. It was further proposed that if these features were incorporated into a cyclic structure with restricted conformations this would be beneficial entropically. The DuPont Merck group further reasoned that these effects would reduce the need for multiple interactions in the specificity pockets and permit the design of smaller inhibitors.

A pharmacaphore search of a 3D database suggested using a six-membered ring to position a flap water mimic. After a few rounds of design, the seven-membered cyclic ureas were invented. (See Figure AP*34.) The seven-membered ring was design to incorporate the diol functionality, to interact with the catalytic aspartates; the urea functionality was introduced into the design for synthetic considerations and the potential of the urea oxygen to make stronger hydrogen bonds to the flap. Model building experiments suggested the urea nitrogens as handles for introducing groups into the S_2 subsites. The predicted optimal stereochemistry for cyclic ureas with substituents on nitrogen was 4*R*, 5*S*, 6*S*, 7*R* that is derived from the unnatural D-phenylalanine. Thus, relatively small, rigid, C_2-symmetrical HIVp inhibitors were incrementally designed. Some of these compounds proved to be potent enzyme inhibitors and anti viral agents. As can be seen, one of these compounds (**AP-27a**) is a potent, low molecular weight inhibitor. A crystal structure between **AP-27b** and HIVp has been reported (PDB entry: 1HVR).[107]

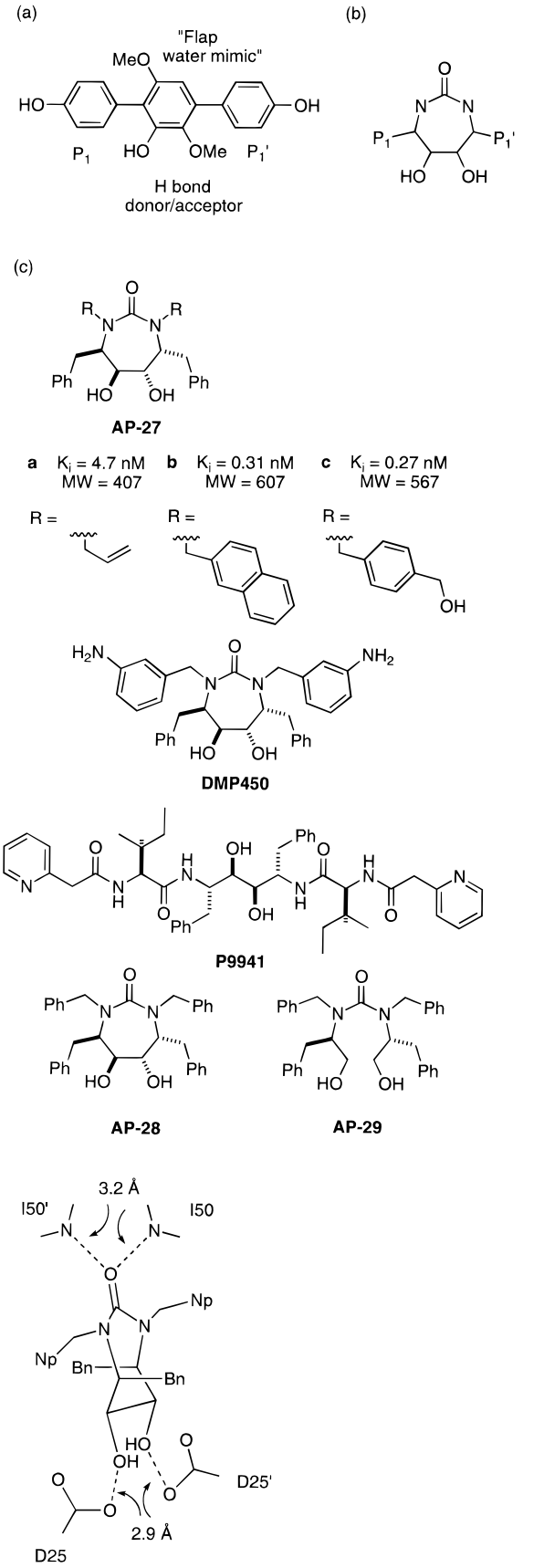


Figure AP*34. (a) Hit from a pharmacaphore search of a 3D database. (b) Designed cyclic urea ring system incorporates elements of "flap" water into the inhibitor. (c) Structures of various inhibitors. (d) Schematic of interactions of the cyclic urea ring system with HIVp. The urea oxygen replaces the "flap" water found in most HIVp inhibitor complexes.

Inhibitor **AP-27b** binds to HIVp in a completely symmetrical fashion. The *S,S* diol group binds in a fashion that is identical to that observed for **A-76928**, in which each hydroxyl group makes nearly equivalent interactions with a single aspartate. The structure shows that the urea carbonyl has displaced the flap water by making hydrogen bonds to the backbone NH's of Ile-50 and Ile-50′. The phenyl groups each make hydrophobic contacts with an S_1 subsite and the napthyl groups each make hydrophobic contacts with an S_2 subsite. The compound bound as modeled. The combination of making a cyclic compound and displacing the flap water has resulted in a series of low molecular weight HIVp inhibitors. For comparison, an acyclic analog, **P9941**, which does not displace the flap water has a $K_i = 9$ nM measured under the same assay conditions. The invention of these cyclic ureas is a classic in protein structure based design. The symmetry of these compounds is coincident with the symmetry of the protein. Displacement of the flap water targets a structural feature only present in retroviral aspartic proteases.

On the basis of pharmacokinetic and safety factors, **DMP323** (**AP-27c**) was chosen as a clinical candidate. That clinical trial was terminated due to highly variable human oral bioavailability. Compound **DMP450** is presently in human clinical trials.[108]

Cyclic urea **AP-27a** (MW = 407) is twice as potent as the acyclic compound **P9941** (MW = 764). The enhanced potency with reduced size can be attributed to a combination of two factors, displacement of the flap water and preorganized substituents on the seven-membered ring. The entropic cost of restricting a rotor, within a hydrocarbon chain, has been estimated as -0.38 to -0.86 kcal mol^{-1} ($T\Delta S$ at 300 K). This estimate was derived from consideration of the entropy of fusion within homologous series of alkanes, alkyl carboxylic acids, and 2-methyl ketones upon crystallization.[7] For the cyclic ureas, five rotatable bonds are restricted, thus, the ring constraint provides at most 1.9−4.3 kcal mol^{-1} stability to the complex. The presence of stereochemically defined substituents in the acyclic core of **P9941** makes these bonds not free rotors. Thus, the entropic benefit of cyclization is probably considerably less that the 1.9−4.3 kcal mol^{-1} maximum. As was previously discussed, displacement of the catalytic water molecule with a hydroxyl group is favorable from an entropic viewpoint. Replacement of the flap water should also be favorable from an enthalpic standpoint since the protein ligand complex does not trap a water molecule from bulk solvent. The catalytic and flap water (this water usually has a low crystallographic temperature factor) molecules can be classified as being tightly bound to HIVp. From entropy estimates it has been suggested that the free energy cost of trapping a tightly bound water molecule is close to 2 kcal mmol^{-1} at 300 K. Thus, replacement of the flap water, like the catalytic water, probably contributes considerably to the stability of the complex.

The DuPont Merck group has reported additional experiments which give further insight into these compounds.[109] The enantiomer of **AP-27c** was prepared and found to be a much weaker inhibitor of HIVp ($K_i = 1650$ nM). The 4900-fold difference in K_i between enantiomers confirmed earlier modeling predictions that the *RSSR* stereochemistry in **AP-27** is the preferred one. A small-molecule crystal structure of **AP-27c** was obtained and compared to its bound conformation. The two structures were very similar, suggesting that **AP-27c** is highly preorganized for binding. Extensive NMR studies of **DMP450** in water also indicate that the ring conformation is preorganized for binding. Cyclic urea **AP-28** and the corresponding seco analog **AP-29** were compared with each other. Cyclic urea **AP-28** has $K_i = 2.5$ nM and the acyclic urea **AP-29** has $K_i = 6700$ nM. This corresponds to a 4.8 kcal mol^{-1} difference in binding energy. On the assumption that **AP-29** binds to HIVp in a loose, relatively unstrained "pseudo cyclic" conformation close to that of **AP-28**, then much of the 4.8 kcal mol^{-1} difference between the cyclic and acyclic compound can be attributed to the conformational preorganization of **AP-28**.

Since no crystal structure of **AP-29** has been obtained, the 4.8 kcal mol^{-1} energetic value may or may not be accurate.

NMR studies[110] have been performed for the complexes of HIVp with both **DMP323** (**AP-27c**) and **P9941**. For **DMP323**, it was proposed that the catalytic aspartates (0,0 protonation state) and the diol groups form two hydrogen bonding networks in dynamic equilibrium. For both complexes, the two protein monomers are magnetically equivalent, demonstrating that the average dimer conformer is symmetric on the NMR time scale when bound to each symmetric inhibitor. Hydrogen-bond formation to either the flap water or the urea carbonyl of **DMP323** requires some asymmetry at the tips of the flap (Ile-50 and Gly-51). It was proposed that a conformational change between two isoenergetic, asymmetric flap conformations takes place on the time scale of $\sim$10 μs. This conformational change makes the enzyme symmetrical on the NMR time scale. For **P9941** it was proposed that the lifetime of an individual flap water molecule is between 9 ns and $\sim$100 ms. The NMR studies reveal a dynamic process occurring during ligand complexation that is not evident from crystallographic studies.

As previously discussed, NMR studies provided evidence that the catalytic aspartates of HIVp have different ionization states in complex with the asymmetric inhibitor **KNI-272**.[74]

Since the disclosure of these cyclic ureas other groups have reported on the design of conceptually related inhibitors.[111−115]

Two reports have independently described the discovery, by random screening, of 4-hydroxycoumarins[116] and 4-hydroxy-2-pyrones[117] as competitive inhibitors of HIVp. Compound **AP-30** had $K_i = 1$ μM and compound **AP-31** had $K_i = 0.7$ μM. Crystal structures of the complexes of these compounds with HIVp show that the hydroxyl group of these compounds interacts with the catalytic aspartates and the coumarin lactone replaces the flap water. (See Figure AP*35.) It is interesting to note the similarity between the 3D-database computer "hit" found by the DuPont-Merck group and the 4-hydroxycoumarin and 4-hydroxy-2-pyrone hits found by screening com-

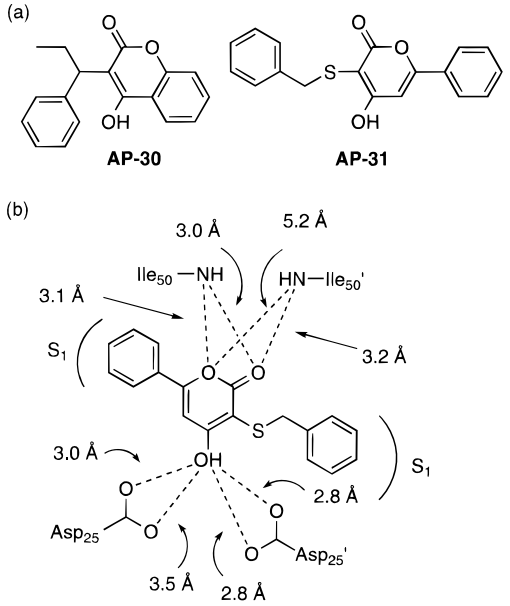


Figure AP*35. (a) Structure of screening lead inhibitors that displace the "flap" water. (b) Schematic of interactions of **AP-31** with HIVp. The lactone group replaces the "flap" water.

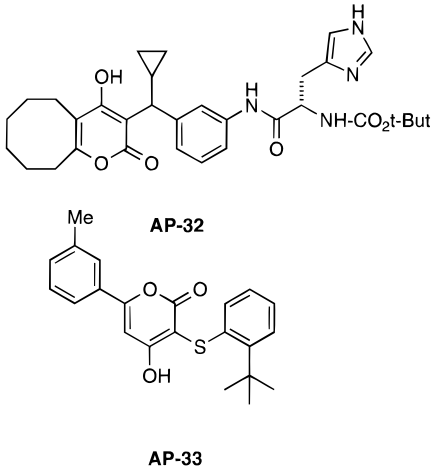


Figure AP*36. Structures of pyridone-based HIVp inhibitors.

pound libraries. The discovery of these compounds, and the structures of their complexes with HIVp have provided a new starting point for design.[116,118]

Structure-assisted design based on these screening leads, and the structure of their complex with HIVp has led to the preparation of potent, nonpeptidyl HIVp inhibitors. Two examples of potent compounds resulting from these studies are **AP-32** ($K_i = 3$ nM)[119] and **AP-33** ($K_i = 3$ nM).[118] (See Figure AP*36.) Compound **AP-33** is achiral, has a molecular weight of 366 and is the smallest potent HIVp inhibitor reported to date.

In another study which sought to optimize the potency of **AP-31**, the more potent analog **AP-34** ($K_i = 34$ nM) was reported. The crystal structure of **AP-34** bound to HIVp was determined.[120] Comparison of this structure with the related **AP-31** revealed a surprise. Similar key interactions were observed for the two inhibitors, hydrogen bonding to catalytic aspartates and displacement of the flap water by the lactone. However, the pyran-2-one rings of **AP-31**

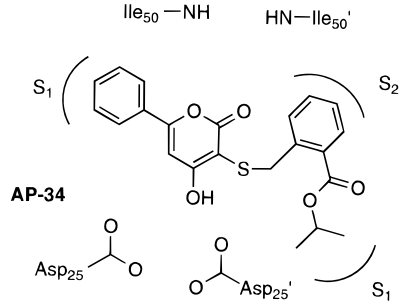


Figure AP*37. The structure of the complex between **AP-34** and HIVp revealed an inversion in binding mode.

and **AP-34** bind at different angles in the enzyme. This results in the sulfurs being oriented 1.3 Å apart in the two complexes. For the **AP-34** complex, the 6-phenyl ring binds in an S_1 subsite, however, the other phenyl ring now occupies an S_2 subsite and the isopropyl group occupies the other S_1 subsite. Thus, for **AP-34** an inversion of binding modes between S_1 and S_2 binding sites was observed. (See Figure AP*37.)

D. Summary and Perspective: Lessons from HIV-1 Protease

From the many structures of HIVp inhibitor complexes with different structural frameworks and different potencies, one common feature is that most of the potent compounds make electrostatic interactions with the catalytic aspartates and displace the catalytic water molecule. As previously discussed, displacement of the catalytic water is highly favorable from both an enthalpic and an entropic standpoint. This general principle was known from the beginning of HIVp inhibitor research. The studies of pepstatin and related analogs clearly laid the groundwork for the initial design of HIVp inhibitors. Upon carefully looking at the details of these interactions, several different binding modes were observed. Perhaps the most surprising was the observed change in stereochemical preference at the carbon bearing the hydroxyl group going from **JG-365** to **Ro 31-8959**. This observation was in contrast to the general principles that were derived from inhibition of mammalian and fungal aspartic proteases. What has emerged from examining the structures of varied HIVp inhibitors and examination of their interactions with the catalytic aspartates is a new general principle. Potent aspartic protease inhibitors result from a combination of favorable electrostatic interactions with the catalytic aspartates (displacing the water) and favorable protein−ligand interactions in the flanking subsites with minimal strain in the linking groups. On the basis of some of the unexpected results observed, it would make sense, from a design standpoint to individually prepare and evaluate each stereoisomer at the carbon bearing the hydroxyl group designed to interact with the catalytic aspartates.

Another observation was that in several cases inhibitors bound to the protein in a different conformation than originally modeled. These results clearly demonstrate the significance of obtaining cocrystal structures as a prelude to further model-based design. In fact, many of the clinical candidates are

based upon the (*R*)-hydroxyethylamine isostere first observed by the Roche group.

In some cases large conformational changes in the protein were observed in response to large groups of the ligand. These changes demonstrated that certain regions of the protein were quite flexible. This suggests that during the design process attempts should be made to probe protein flexibility using groups that appear to be slightly too large.

One important problem solved in the development of clinically useful HIVp inhibitors was the preparation of orally bioavailable aspartic protease inhibitors. This required making inhibitors of lower molecular weight and less peptidic character. During these investigations several novel peptidomimetics were designed. It is instructive to compare and contrast some of these inhibitors with peptidic compounds. The work of the SmithKline group demonstrated that potent enzyme inhibition and submicromolar antiviral activity could be observed in a compound that spans from P_2 to P_2' (MW = 510). While the cocrystal structure of this compound has not been reported, the cocrystal structure of a related imidazole-based compound has been reported. For the following discussion, a model of **AP-15** was prepared by substituting the imidazole group in PDB entry 1SBG with a $CONH_2$ group. The modeled complex of **AP-15** buries 1280 Å^2 of ligand (835 Å^2, 729 Å^2 of nonpolar surface area) and protein (445 Å^2, 236 Å^2 of nonpolar surface area) upon complex formation. In this structure strong hydrogen bonds are observed with the flap water and strong hydrogen bonds are made with the backbone NH's of Gly-27 and Gly-27′. For **AP-15** and peptidyl inhibitors both hydrophobic and substrate-like hydrogen bonding contribute to the stability of the complex.

The compound **Ro 31-8959** is clinically useful primarily because it is a very potent antiviral agent. The complex buries 1485 Å^2 of ligand (907 Å^2, 750 Å^2 of nonpolar surface area) and protein (539 Å^2, 324 Å^2 of nonpolar surface area) surface area. Comparing **Ro 31-8959** with **AP-15** shows that the P_1 site is similar, phenyl groups in S_1 subsite, hydrogen bond formation with Gly-27 and to the flap water. The P_2 and P_3 subsites are very different. While the *tert*-butoxy group of **AP-15** only makes hydrophobic contacts in the S_2 subsite, the asparagine side chain of **Ro 31-8959** makes much less VDW contacts with the protein. As previously discussed the amide group of the asparagine side chain makes two hydrogen bonds to the protein, a long (3.4 Å) hydrogen bond between the CO group and backbone NH of Asp-30 and a strong hydrogen bond between the NH_2 group and an anti-lone pair of the side chain of Asp-30. Thus, favorable interactions in the S_2 subsite can be made either from hydrophobic contacts or by favorable electrostatic contacts. **Ro 31-8959** has a large quinoline ring that makes favorable contact with the P_1 phenyl group. The quinoline ring and the adjacent carbonyl are in a low-energy local conformation. It is quite possible that these two groups interact with each other in solution and induce a bioactive conformation by ligand hydrophobic collapse. On the prime side of the inhibitors, the DIQ group makes more extensive hydrophobic contact with the protein than does the phenyl group of **AP-15**. As previously discussed the protein undergoes a conformational change to accommodate the large DIQ group. The *tert*-butyl group of **Ro 31-8959** make hydrophobic contact with an S_2 subsite in a similar fashion as the isopropyl groups of **AP-15**. The amide group of **Ro 31-8959** does not interact with Gly-27′ but rather with structural waters. **Ro 31-8959** is a very potent inhibitor by making extensive protein−ligand contacts, both hydrophobic and electrostatic. The molecular weight of **Ro 31-8959** is 670, much larger than **AP-15** with a MW = 510. The clinical utility of **Ro 31-8959** derives from its high potency. Comparison of it with **AP-15** shows it uses its larger size to make more extensive protein−ligand contacts than does **AP-15**. The P_3 to P_1 groups form substrate-like hydrogen bonds with the protein. It will now be instructive to compare **Ro 31-8959**, with a related, but smaller inhibitor, **AG-1343**.

Compound **AG1343** is a potent HIVp inhibitor with a MW = 567 with high structural similarity to **Ro 31-8959**. The complex buries 1271 Å^2 of ligand (813 Å^2, 707 Å^2 of nonpolar surface area) and protein (460 Å^2, 260 Å^2 of nonpolar surface area) surface area, this is approximately the same as for **AP-15**. Compared to **Ro 31-8959**, **AG1343** is slightly less potent as an enzyme inhibitor (by a factor of 2.2 under the same assay conditions) and a weaker antiviral agent (by a factor of ~4); however, the molecular weight and log *P* are decreased and the aqueous solubility was enhanced. These altered physical properties presumably are responsible for increased oral bioavailability.[80] While the clinical utility of **Ro 31-8959** is due primarily to its high antiviral potency, the clinical utility of **AG1343** is due to a combination of antiviral potency and high oral bioavailability. **AG1343** and **Ro 31-8959** have identical P_1' and P_2' groups and these bind to HIVp in an identical fashion. They differ in that **AG1343** has a much smaller, nonpeptidic P_1 and P_2 groups. While the *tert*-butyl carbamate of **AP-15** makes hydrophobic contact in an S_2 subsite and the asparagine side chain of **Ro 31-8959** makes favorable electrostatic contacts in that subsite, the 2-methyl-3-hydroxybenzamide ring of **AG1343** simultaneously makes favorable hydrophobic and electrostatic contacts in that site. In addition to the methyl group of that ring making favorable hydrophobic contact with the protein, it also serves to drive the carbonyl group out of plane, thus allowing for hydrogen-bond formation with the flap water. The phenol group of that ring makes a favorable hydrogen bond to an anti-lone pair of Asp-30 (2.9 Å) in much the same manner as the NH_2 group of the asparagine of **Ro 31-8959**. The *S*-phenyl group in **AG1343** makes extensive hydrophobic contact with an S_1-S_3 subsite and causes a conformational change in the protein. The P_1 and P_2 aromatic rings make an edge-to-face interaction in the complex. This appears to be a low energy local conformation. It is possible that these two groups also interact with each other in solution and induce a bioactive conformation by ligand hydrophobic collapse. Compound **AG1343** has three fewer rotatable bonds compared to **Ro 31-8959**. The favorable interactions made by the P_1 and P_2 groups of **AG1343**

allow for a reduction in inhibitor size, a decrease in peptidic character, and improved pharmacological properties compared to **Ro 31-8959**.

Compound **VX-478** is a low molecular weight (506), potent, orally bioavailable HIVp inhibitor. The complex buries 1223 of ligand (814 Å^2, 659 Å^2 of nonpolar surface area) and protein (410 Å^2, 220 Å^2 of nonpolar surface area) surface area, this is slightly less than **AP-15**. Compound **VX-478** has the same central (*R*)-hydroxyethylamine isostere as does **Ro 31-8959** and **AG1343**; however, it has different P_2, P_1', and P_2' groups as well as sulfonamide moiety on the central nitrogen. In the P_2 position it has a 3(*S*)-tetrahydrofuryloxy group. Unlike the *tert*-butoxy group in **AP-15** that only provides hydrophobic contact with the protein, this group also forms weakly polar interactions with the backbone NH's of Asp-29 and Asp-30 (3.4 and 3.5 Å, respectively). The conformation of the THF ring is in an energetically preferred axial orientation. The Merck group had previously shown in a hydroxyethylamine series that a 3(*S*)-tetrahydrofuryloxy group is more than 15 times more potent than a *tert*-butoxy group in the P_2 position. Compound **VX-478** does not form a hydrogen bond to either Gly-27 or Gly-27′. One of the sulfonamide oxygens interacts with the flap water, while the other is partially buried in a hydrophobic environment and 3.4 Å from the flap water. The sulfonamide group appears to be in a local conformational minima. The P_2' aromatic makes hydrophobic contact with an S_2 subsite, and the amino substituent makes a hydrogen bond to Asp-30′. Thus, the P_2' group, like the P_2 group makes favorable hydrophobic and electrostatic interactions with its adjacent S_2 subsite. Compound **VX-478** has a relatively small hydrophobic P_1' group. This results in the two S_1 subsites being of the same size and shape. Unlike **Ro 31-8959** and **AG1343**, no conformational change in these subsites is induced by a large group. The favorable protein−ligand interactions made by **VX-478** allowed for a reduction in inhibitor size, a decrease in peptidic character, and improved pharmacological properties compared to **Ro 31-8959**.

In contrast to substrate like inhibitors, neither **AG1343** nor **VX-478** make a direct hydrogen bond to the backbone NH of either Asp-29 or Asp-29′. For **AG1343** the face of the phenol ring lies above Asp-29 NH and the Asp-29′ NH is solvent exposed. For **VX-478** the THF ring makes a weak electrostatic interaction with Asp-29 NH, and Asp-29′ NH (close to the sulfonamide aromatic ring) is solvent exposed. Like **AP-15**, both of these nonpeptidic compounds form favorable hydrophobic interactions with HIVp and favorable electrostatic interactions with the catalytic aspartates and the flap water. Compounds **AG1343** and **VX-478** make favorable electrostatic in the vicinity of the S_2 subsites, but do it in a manner unlike **AP-15** and peptidic inhibitors. These two compounds represent the only linear acyclic compounds, with clinical potential, with molecular weights under 600. One feature these compounds seem to have in common is that most of the binding groups appear to be in low-energy conformations and are somewhat preorganized for binding.

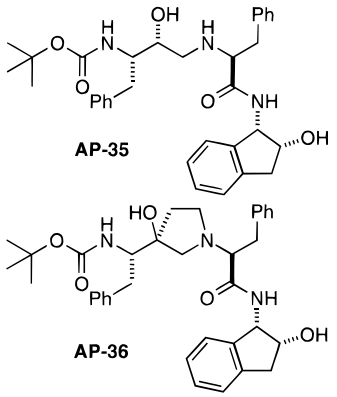


Figure AP*38. Structures of HIVp inhibitors, **AP-36**, a conformationally constrained analog of **AP-35**, is significantly less potent than **AP-35**.

The cyclic urea **DMP323** and the 4-hydroxypyran-2-one **AP-33** are two low molecular weight, low nanomolar HIVp inhibitors. The high potencies and low molecular weight of these compounds are primarily due to a combination of two factors, displacement of the flap water and conformationally restricted ring systems. The water displacement process and restricted conformational space result in a high entropic component to the binding phenomenon.

The high potency of these compounds may be partly due to the inherent stability of extended conformations of these compounds in aqueous solution.[118] Thus, these ring systems may prevent unproductive hydrophobic collapse that can be detrimental to high-affinity binding. In the case of **Ro 31-8959** the P_1 phenyl and P_3 quinoline rings interact with each other in the complex with HIVp; it is possible that a similar aggregation also occurs in aqueous solution. If so, this would be an example of productive hydrophobic collapse. Likewise, the aromatic−aromatic interaction observed for **AG1343** may also represent productive hydrophobic collapse.

The Merck group[121] reported that the hydroxyethylamine-based inhibitor **AP-35** had an $IC_{50} = 250$ nM. (See Figure AP*38.) In an effort to improve the potency of this series, by enforcing the bioactive conformer, they designed and prepared a number of conformationally constrained compounds based on **AP-35**. Inhibitor **AP-36** was prepared and was found to be less potent with an $IC_{50} = 600$ nM. The crystal structure of **AP-36** has been reported. As expected, the two benzyl groups are each in an S_1 subsite. However, several unexpected features were observed: (1) The inhibitor only had an occupancy of 0.5. (2) The geometry of the catalytic aspartates is very strained, probably as a result of steric crowding caused by the ring system of the inhibitor. Additionally, the central hydroxyl group does not appear to be in good contact with the aspartates. (3) The conformation of the P_2 groups was considerably skewed to the extent that the carbonyl oxygen does not appear to interact with the flap water (O−O distance = 5.7 Å). These unusual features are likely responsible for the poor potency of these compounds.

Another study by the Merck group[122] on conformationally constrained HIVp inhibitors was also reported. The hydroxyethylene isostere **AP-37** (Figure AP*39) has an $IC_{50} = 0.3$ nM. The conformationally

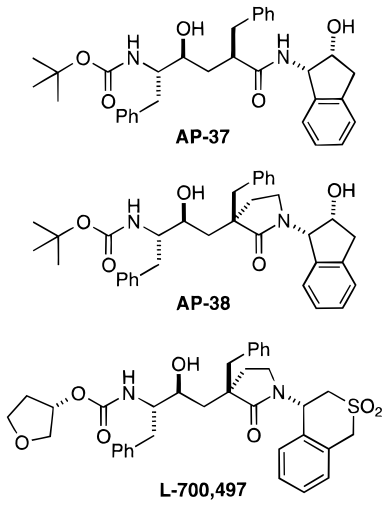


Figure AP*39. Structures of HIVp inhibitors, **AP-38**, a conformationally constrained analog of **AP-37**, is >2 orders of magnitude less potent than **AP-37**.

constrained analog **AP-38** was prepared and determined to be more than 2 orders of magnitude less potent with an $IC_{50} = 37$ nM. Optimization of the P_2 and P_2' groups of **AP-38** resulted in **L-700,497** having an $IC_{50} = 1.8$ nM. The cocrystal structure of **L-700,497** with HIVp was reported. The central hydroxyl group is situated between the catalytic aspartates, the flap water is present and hydrogen bonds to the two central carbonyls, and the benzyl groups each occupy an S_1 subsite. The novel P_2' isothiabenzopyran group is in an S_2 subsite and one of the sulfone oxygens is within hydrogen bonding distance of the backbone NH of Asp-29 (3.2 Å). Comparison of the complex structures of **L-700,497** and a close analog of **AP-37** shows that the major difference is that for the acyclic compound the indanamide NH makes a hydrogen bond to the backbone carbonyl of Gly-27 while the lactam cannot make this interaction. This results in that carbonyl being displaced 1 Å relative to the acyclic compound. In addition, the lactam ring appears to experience nonbonded interactions that change the location of the P_1' carbonyl group relative to the acyclic compound. Thus, rather subtle changes in structure are responsible for the observed significant loss of potency.

Two other groups[123,124] have reported on macrocyclic systems in which the P_1 and P_3 groups are connected to each other. In these cases reduced potency was reported relative to the acyclic analogs. The Queensland group[124] reported the preparation of a macrocyclic analog of **JG-365** (**AP-39**, $K_i = 39$ nM; PDB entry: 1CPI (in the same assay the *S* isomer of **JG-365** had $K_i = 1$nM)). (See Figure AP*40.) The cocrystal structure of **AP-39** with HIVp was reported and the bound conformation was very similar to that observed for **JG-365**. However, the constrained macrocycle shows a 40-fold loss in potency compared to the acyclic analog. One possible explanation for the loss in potency is that the constrained analog is predominantly constrained in conformations that are not related to the bioactive conformation.

An informative discussion[125] of the consequences of ligand conformation in solution on complex stability has been published. In this analysis (eq 1) the observed K_d will be a function of the intrinsic K_d of the bioactive conformer(s) and its population in solution K_{eq}^{AC}. As defined in eq 2, K_{eq}^{AC} is the equilibrium constant between ensembles of bioactive conformers, C_A, and inactive (unproductive) conformers, c_n, in aqueous solution. Several consequences for ligand binding can be inferred from these relationships. The observed K_d will equal the intrinsic K_d when only bioactive conformers are present in solution ($K_{eq}^{AC} = 0$). As the population of bioactive conformers in solution decreases an energetic penalty must be paid and the observed K_d is reduced relative to the intrinsic K_d of the bound conformer. That is, even though the ligand may make many productive interactions with the protein, resulting in a tight intrinsic K_d, the population of nonproductive conformers in solution makes the observed K_d less tight. Let us now consider two extreme cases. In case 1 there are 100 conformers in solution of equal energy, only one of these is a bioactive conformer, $K_{eq}^{AC} = 99$. In case 2 there are only two conformers in solution that differ in energy by 2.7 kcal mol^{-1}. The less stable conformer is the bioactive one, $K_{eq}^{AC} = 99$. Both of these situations are equally detrimental to tight ligand binding.

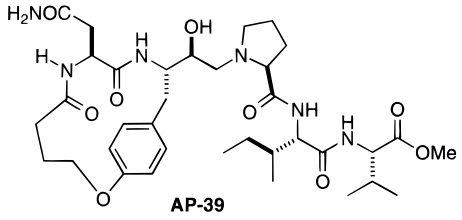


Figure AP*40. AP-39, a macrocyclic analog of **JG-365**, is less potent than **JG-365**. The structure of the **AP-39**-HIVp complex shows that it binds like **JG-365**. It is likely that the poorer inhibition of **AP-39** is due to the macrocyclic constraints disfavoring the bound conformation.

$$K_d^{obs} = K_d^{int}[1 + K_{eq}^{AC}] \quad (1)$$

$$K_{eq}^{AC} = \frac{\{c_1^i, c_2^i, ..., c_n^i\}}{\{C_A\}} \quad (2)$$

For the case of **JG-365** and the macrocyclic analog **AP-39** described above the observed cocrystal structures show very similar interactions, so it can be assumed that the intrinsic K_d for both ligands are approximately equal. However, the constrained macrocyclic compound is 40-fold less potent. It is possible that the reduced potency for the macrocyclic analog is due to a ligand conformational effect. For the acyclic compound it is likely that it has a large number of conformers in solution, some of them bioactive and some unproductive (similar to case 1 described above). For the macrocycle it is likely that it has fewer conformations in solution. The lower observed K_i for the macrocyclic compound is consistent with being about 40 times higher for the macrocycle. For this to be the case the bioactive conformer must be of high energy in solution (similar to case 2 described above). On the basis of this analysis, an explanation for the loss in potency is that the

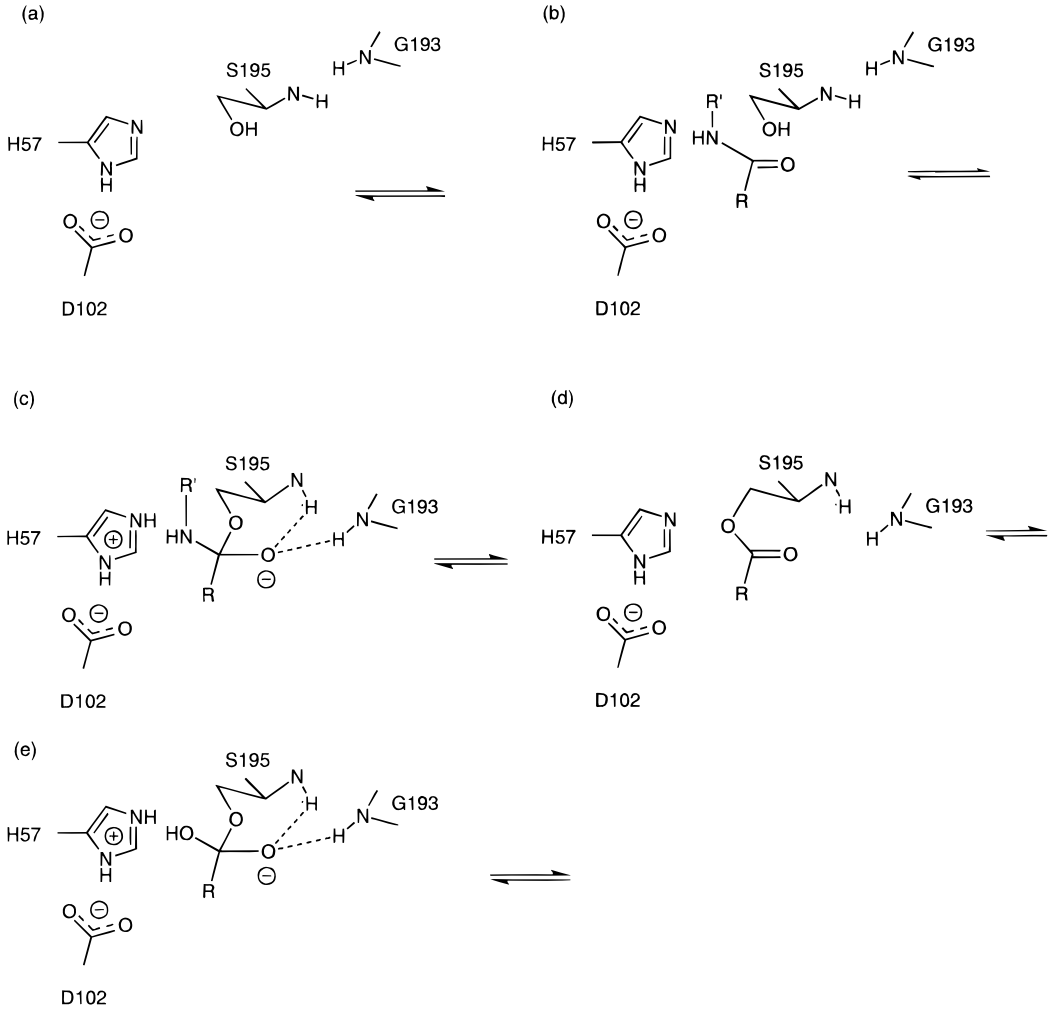


Figure SP*1. Mechanism of amide hydrolysis catalyzed by serine proteases; see text for details.

constrained analog is predominantly constrained in conformations that are not related to the bioactive conformation.

These studies on constrained inhibitors of HIVp clearly demonstrate that restricting the conformation space of an inhibitor can be beneficial to binding when the conformation is biased toward the bioactive conformer. However, if the parts of the constricting group make unfavorable interactions with the protein, or bias an nonbioactive conformer then deleterious effects on binding can be obtained.

IV. Inhibition of Serine Proteases

A. Introduction

Serine proteases have been one of the most extensively studied family of enzymes. There are several different classes of serine proteases which have been examined to date. Of these, the trypsin/chymotrypsin class has been the most thoroughly characterized.[126] In terms of drug design, the most extensively investigated protein targets in this class have been human thrombin and human neutrophil elastase. The assemblin class of herpes proteases[127−131] and hepatitis C protease[132,133] appears to be targets which will receive much attention in the future.

The proposed mechanism of amide hydrolysis by serine proteases is shown in Figure SP*1.[126] The active site in the enzyme resting state consists of a "catalytic triad" and an oxyanion hole (a). The groups of the catalytic triad are Ser-195, the nucleophile, His-57, the general base and general acid, and Asp-102, a group that properly orients His-57. The oxyanion hole consists of the backbone amide NH's of Ser-195 and Gly-193. A substrate binds to the enzyme to form the noncovalent Michaelis complex (b). The hydroxyl group of Ser-195, under general base catalysis by His-57, attacks the carbonyl carbon of the substrate to form the tetrahedral intermediate (c). The oxyanion of the tetrahedral intermediate is stabilized by forming hydrogen bonds to the oxyanion hole. Proton transfer from His-57, acting as a general acid, to the amine of the tetrahedral intermediate occurs and the product amine (R'-NH_2) leaves as the tetrahedral intermediate breaks down to the covalent acyl enzyme complex (d). The acyl enzyme is attacked by water, assisted by general base catalysis by His-57, to generate a new tetrahedral intermediate (e). Breakdown of this intermediate, assisted by general acid catalysis by His-57, results is the product acid (R-CO_2H) in which Ser-195 is the leaving group. (e) For chymotrypsin-catalyzed hydrolysis of benzoyl-L-arginine amide the $^{14}N/^{15}N$ kinetic isotope observed was close to unity. This was considered as evidence that there is very little C−N bond cleavage occurring in the transition state.[134] The

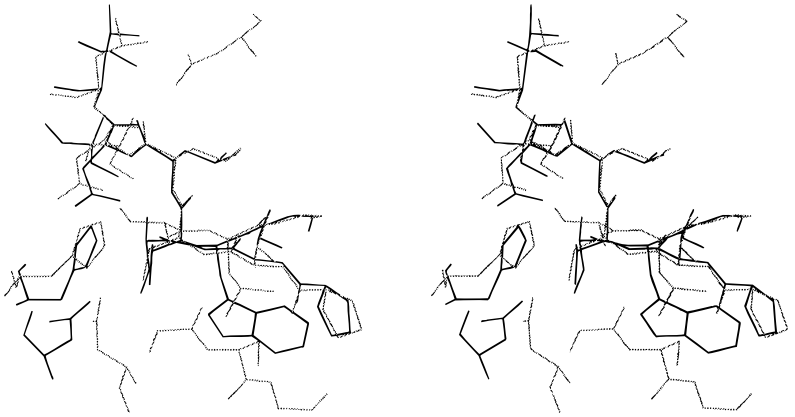


Figure SP*2. Stereoview of overlap of the active sites of chymotrypsin (bold) and subtilisin (gray) showing common inhibitor eglin-C. Overlap was performed by aligning the backbones of the common inhibitor eglin-C from residues I-41 to I-46. See text for details.

kinetic evidence suggests that the transition state catalyzed by chymotrypsin is close to, and presumably structurally and electronically similar to, the tetrahedral intermediate.

Convergent evolution has produced structurally distinct families of serine proteases with similar catalytic mechanisms. Two of the best studied groups are the mammalian chymotrypsin, which includes thrombin and elastase, and bacterial subtilisin superfamilies of homologous enzymes.[135] The chymotrypsin family of enzymes has two antiparallel β-barrel domains while the subtilisin family is of the α/β type. While the protein architecture of these two families of enzymes are quite different, the details of their active sites are very similar. The active-site atoms that comprise the catalytic triad and oxyanion hole are nearly identical in these two enzyme classes.

One type of naturally occurring inhibitors of serine proteases are proteinaceous inhibitors. In fact, there are many structural and mechanistic classes of these natural protein serine protease inhibitors.[136] Eglin-C, a member of the potato inhibitor family 1 (PI-1) class of inhibitors, is quite informative since it is a substrate-based inhibitor and it inhibits both α-chymotrypsin (PDB entry: 1ACB)[137] and subtilisin carlsberg (PDB entry: 1CSE).[138] High-resolution crystal structures for both of these complexes have been determined and illustrate the similarities and differences between two structurally distinct but mechanistically similar serine proteases. These proteinaceous inhibitors are substrate-like active-site inhibitors.

The structures of α-chymotrypsin and subtilisin carlsberg in complex with eglin C were superimposed by aligning the backbones of the common inhibitor from residues I-41 to I-46. Figure SP*2 shows the active sites of these two structurally distinct enzymes aligned by this superpositioning. From this alignment the active-site nucleophiles (serine OH's), and general acid/base (histidine) are closely superimposed by this alignment.

In the eglin C−chymotrypsin complex, the inhibitor is bound in a substrate-like fashion. The carbonyl carbon of Leu-I-45 is in "sub-van der Waals" contact (2.8 Å) with the oxygen of Ser-195, the carbonyl oxygen is in the oxyanion hole-forming hydrogen bonds to the backbone NH's of Ser-195 and Gly-193

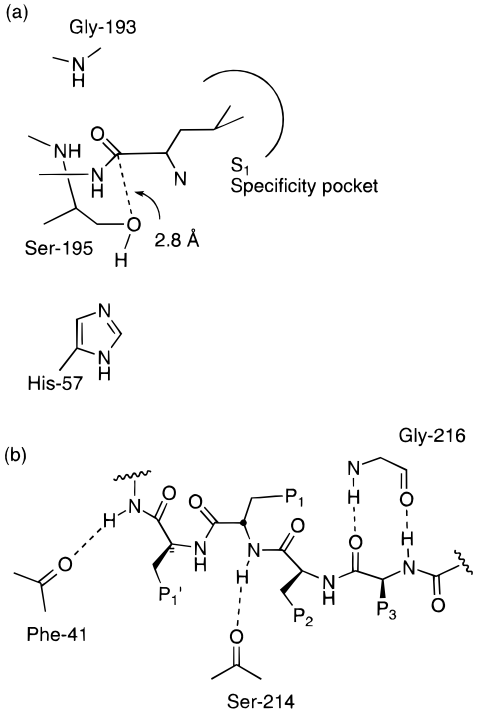


Figure SP*3. Eglin C−chymotrypsin complex. (a) Interaction of P_1-P_1' amide of eglin C with chymotrypsin. The leucine side chain makes VDW contacts in the S_1 specificity pocket and the carbonyl carbon is 2.8 Å from the hydroxyl oxygen of Ser-195. (b) The P_3 backbone atoms bind to the protein as an antiparallel β-strand to the main chain of Gly-216 and a long hydrogen bond between the backbone carbonyl oxygen of Ser-214 and the backbone NH of the P_1 residue is present.

with distances of 3.0 and 2.6 Å, respectively. The leucine side chain makes VDW contacts in the large S_1 specificity pocket. This complex resembles a Michaelis complex. (See Figure SP*3.)

Other substrate binding interactions are apparent in this structure. The segment amino terminal to the scissile amide bond (P_3 backbone atoms) binds to the protein as an antiparallel β-strand to the main chain of Gly-216. A long hydrogen bond (3.2 Å) between the backbone carbonyl oxygen of Ser-214 and the backbone NH of the P_1 residue is present. The other backbone atoms from P_2 to P_4 are solvent exposed. On the prime side a hydrogen bond between the P_2' NH and the carbonyl of Phe-41 is present. In

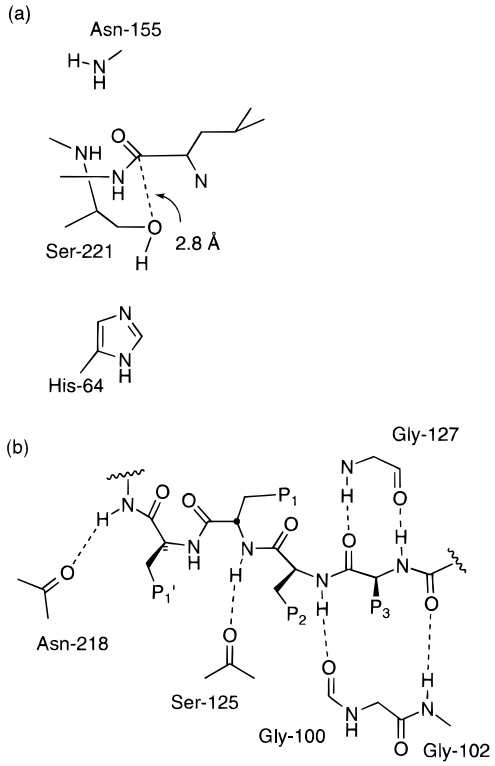


Figure SP*4. Eglin C−substilisn complex. (a) Interaction of P_1-P_1' amide of eglin C with substilisn. The leucine side chain makes VDW contacts in the S_1 pocket and the carbonyl carbon is 2.8 Å from the hydroxyl oxygen of Ser-221. (b) The P_3 backbone atoms bind to the protein as a three stranded β-sheet, making hydrogen bonds to the main chain of Gly-127 on one side and hydrogen bonds to Gly-100 and Gly-102 on the other side, a long hydrogen bond between the backbone carbonyl oxygen of Ser-125 and the backbone NH of the P_2 residue is present.

addition to extensive VDW contacts by the P_1 leucine side chain, the P_2 (threonine), P_4 (proline), P_1' (aspartate), and P_2' (leucine) side chains make VDW contacts with the enzyme. The P_3 (valine) side chain is solvent exposed.

For the eglin C−subtilisin carlsberg complex, the inhibitor is also bound in a substrate-like fashion. The carbonyl carbon of Leu-I-45 is in "sub-van der Waals" contact (2.8 Å) with the oxygen of Ser-221, the carbonyl oxygen is in the oxyanion hole forming hydrogen bonds to the backbone NH of Ser-221 and side chain NH of Asn-155 with distances of 3.0 and 2.7 Å respectively. The amide group between Leu-I-45 and Asp-I-46 has a slight out-of-plane deformation as the carbonyl carbon approaches the nucleophilic serine oxygen (C−O−N−CA improper torsion angle = −165°). This complex thus resembles a distorted Michaelis complex. (See Figure SP*4.)

Other substrate binding interactions are apparent in this structure, and they are different from the chymotrypsin complex. The segment amino terminal to the scissile amide bond (P_3 backbone atoms) binds to the protein to complete a three stranded β-sheet, making hydrogen bonds to the main chain of Gly-127 on one side and hydrogen bonds to Gly-100 and Gly-102 on the other side. A long hydrogen bond (3.8 Å) between the backbone carbonyl oxygen of Ser-125 and the backbone NH of the P_1 residue is present. On the prime side a hydrogen bond between the P_2' NH and the backbone carbonyl of Asn-218 is present. The P_1 (leucine), P_2 (threonine), P_4 (proline), P_1' (aspartate), and P_2' (leucine) side chains make VDW contacts with the enzyme. The P_3 (valine) side chain is solvent exposed.

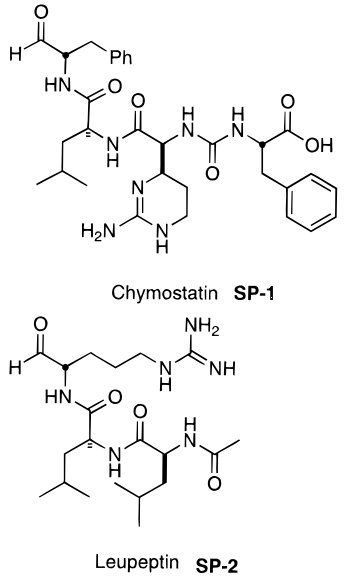


Figure SP*5. Structures of natural product peptide aldehyde inhibitors of serine proteases.

An important issue is why are eglin C, and other proteinaceous inhibitors, which resemble substrates, good inhibitors and poor substrates? For a substrate, binding to the protease active site, to form a Michaelis complex requires that several rotatable bonds must be fixed in the complex. This, of course, is unfavorable from an entropic point of view. The weakly bound Michaelis complex is then converted into the acyl enzyme complex utilizing the ability of the enzyme to stabilize the transition state. The protease binding domain of eglin C is quite rigid, and is preorganized to be complimentary to the protease active site.[136] As a result fewer torsional degrees of freedom need to be frozen out upon complex formation, this difference in entropy loss leads to a deeper energy minimum for the complex.[139]

B. Natural Products as Inhibitors of Serine Proteases

Chymostatin (**SP-1**)[140] and leupeptin (**SP-2**)[141,142] are natural products which inhibit various serine proteases. (See Figure SP*5.) Crystal structures of chymostatin bound to *Streptomyces griseus* protease A (PDB entry: 1SGC)[143] and leupeptin bound to trypsin (PDB entries: 1JRS and 1JRT)[144] have been reported.

Chymostatin was originally reported as an inhibitor of chymotrypsin. *Streptomyces griseus* protease A (SGPA) is a member of the chymotrypsin superfamily, and like chymotrypsin, has a large hydrophobic S_1 specificity pocket. The structure of the chymostatin SGPA complex[143] shows that the aldehyde function has undergone nucleophilic addition by the nucleophilic Ser-195. The remainder of the protein−ligand interactions are similar to those in the eglin C−chymotrypsin complex. The P_1 phenyl ring makes intimate VDW contacts in the large hydrophobic S_1

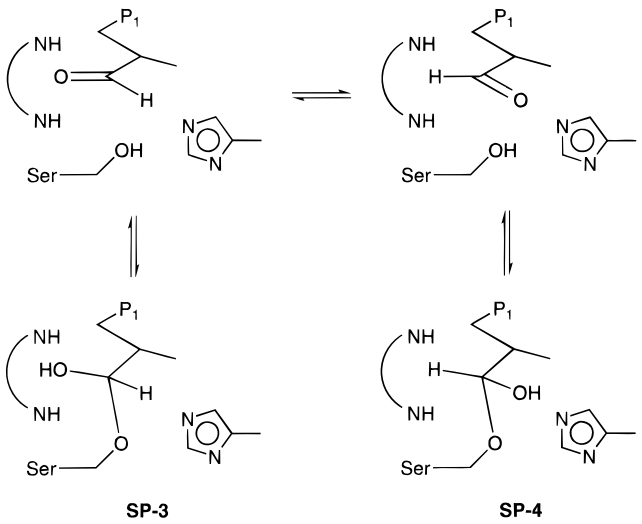


Figure SP*6. Inhibition of serine proteases is a reversible process which often leads to two diastereomeric adducts.

specificity pocket. The P_3 backbone amide atoms bind to the protein as an antiparallel β-strand. The P_2 and P_4 side chains make VDW contacts with the protein while the P_3 side chain and terminal carboxylate are solvent exposed.

Although the aldehyde forms a covalent bond with the protein it is a reversible inhibitor since alcohol addition to aldehydes is a reversible reaction. In the complex there was observed to be two equally populated orientations for the hemiacetal hydroxyl. In one orientation, **SP-3**, the hemiacetal hydroxyl is pointed into the oxyanion hole and interacts with the backbone NH's of Ser-195 and Gly-193. In the other orientation, **SP-4**, the hydroxyl group points toward His-57. It should be noted that these two orientations result from attack on different stereochemical faces of the aldehyde. Compounds **SP-3** and **SP-4** are diastereomers and cannot be interconverted by simple bond rotations. Interconversion between **SP-3** and **SP-4** can only be achieved by reversal to form the aldehyde, bond rotation of the aldehyde, and nucleophilic attack of the other stereochemical face of the aldehyde. (See Figure SP*6.)

The complex between leupeptin and trypsin[144] is very similar. The aldehyde has undergone nucleophilic addition by the active-site serine. In this example the predominant orientation of the hydroxyl group was **SP-4**. During the refinement process, the occupancy of orientation **SP-3** was determined to be 26%. The tendency for peptidyl aldehyde inhibitors to bind in two different orientations with respect to the oxyanion hole suggests that they are not true transition-state analogs.

The major difference between serine protease inhibition by chymostatin and by leupeptin is in the specific enzymes inhibited by each inhibitor. Due to the hydrophobic P_1 side chain of chymostatin, it inhibits enzymes with hydrophobic S_1 specificity pockets. In contrast, leupeptin with an arginine P_1 side chain inhibits serine proteases with complimentary S_1 specificity pockets, such as trypsin, thrombin, and plasmin. Like the S_1 specificity pocket of chymotrypsin, the trypsin S_1 pocket is quite large. However, trypsin has an aspartate residue located deep within the pocket, for chymotrypsin the corresponding residue is a serine. The presence of this aspartate makes trypsin selective for substrates and competitive peptidyl inhibitors, which have a lysine or arginine as their P_1 residue.

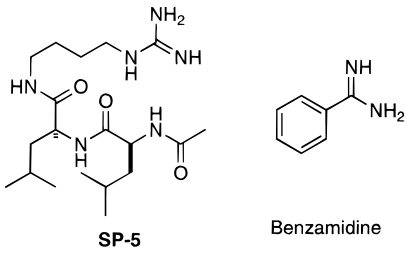


Figure SP*7. Structures of various compounds.

The P_3 side chain of leupeptin stacks against the P_2 side chain and makes VDW contacts with the protein near Trp-215. This requires a different backbone conformation than observed for chymostatin. In the bound conformation no hydrogen bond is formed to the carbonyl of Gly-216, but rather a long (3.86 Å) electrostatic interaction between the acetyl carbonyl and the backbone NH of Gly-219. This results in the carbonyl of Gly-216 being buried and not being satisfied with a hydrogen bond to the ligand. A chymostatin-like conformation would form a hydrogen bond with the Gly-216 carbonyl, but would solvent expose the P_3 leucine side chain.

Leupeptin inhibits trypsin with a K_i = 400 nM. Studies with leupeptin analogs suggest that most of the protein−ligand affinity arises from hemiacetal formation and interactions with the S_1 binding pocket. For example, compound **SP-5** (Figure SP*7), a close analog of leupeptin, lacking only the aldehyde, fails to inhibit trypsin.[144] Another inhibitor, benzamidine, which binds only to the S_1 subsite, but has no rotatable bonds, has K_i = 100 μM[144] (PDB entry: 1BIT).[145]

Compound **A90720A**, a 19-membered cyclic depsipeptide, is a marine natural product which inhibits bovine trypsin (IC_{50} = 10 nM), bovine thrombin (IC_{50} = 275 nM) and human plasmin (IC_{50} = 30 nM) (Figure SP*8). A crystal structure of **A90720A** bound to bovine trypsin has been reported (PDB entry: 1TPS).[146] This inhibitor binds to trypsin in a substrate-like fashion similar to serine protease complexes with "small" protein inhibitors. (See Figure SP*9.)

The arginine side chain occupies the deep, narrow S_1 specificity pocket of trypsin. The guanidine group makes two hydrogen bonds to the side chain of Asp-189, as well as hydrogen bonds to the side chain of Ser-190 and the main-chain carbonyl of Gly-219. In addition, it makes a water-mediated hydrogen bond to Gly-216 and Gly-219 and another to Val-227. The main chain carbonyl carbon is in "sub-VDW" contact (2.8 Å) with the nucleophilic oxygen of Ser-195, the carbonyl oxygen is in the oxyanion hole and makes hydrogen bonds to the backbone NH's of Ser-195 and Gly-193.

In addition to making favorable electrostatic interactions in the S_1 specificity pocket, the inhibitor makes substrate-like hydrogen bonds as an antiparallel β-strand to the main chain of Gly-216. However, unlike a substrate (or protein based inhibitor), it does so with a D-leucine side chain. This unnatural side

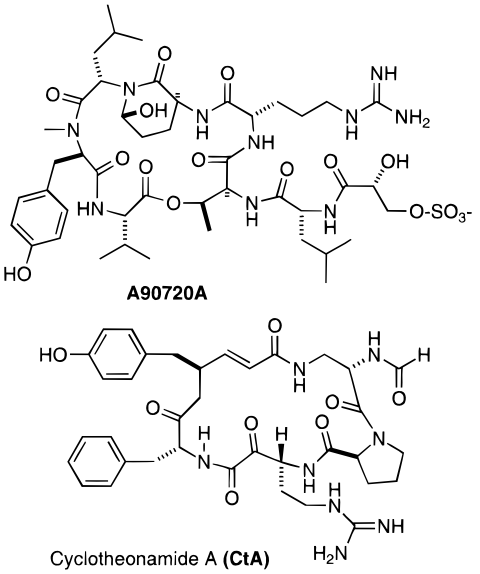


Figure SP*8. Macrocyclic marine natural product inhibitors of trypsin and thrombin.

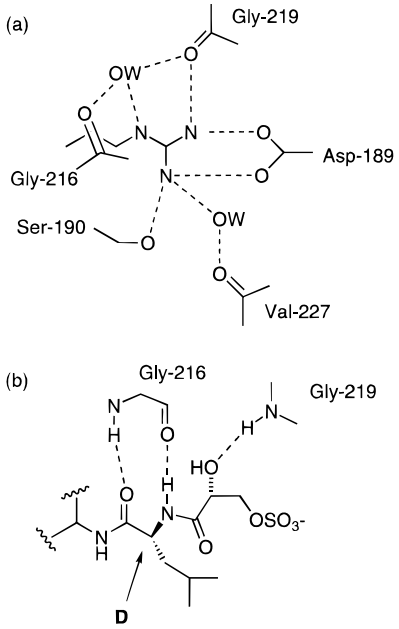


Figure SP*9. (a) Interaction of guanidine side chain of **A90720A** with trypsin S_1 specificity pocket. (b) Hydrogen-bond interactions of **A90720A** with Gly-216 are achieved with a D-leucine side chain.

chain allows the isobutyl group to make favorable hydrophobic contacts with the S_4 subsite and allow the glyceric sulfate to make a favorable, solvent exposed, electrostatic interaction with the protein. The amide NH of the D-leucine side chain is still capable of forming a hydrogen bond to the backbone carbonyl of Gly-216.

The side chains of the macrocycle appear to make nonspecific hydrophobic contacts with the protein surface. Two transannular hydrogen bonds are observed within the macrocyclic ring. These intramolecular hydrogen bonds presumably stabilize the bound conformation of the inhibitor. The 3-amino-6-hydroxy-2-piperidone residue appears to play a key role in rigidifying the macrocycle. Thus, **A90720A** presumably is a potent inhibitor of trypsin because it makes many optimal enzyme−substrate-like in-

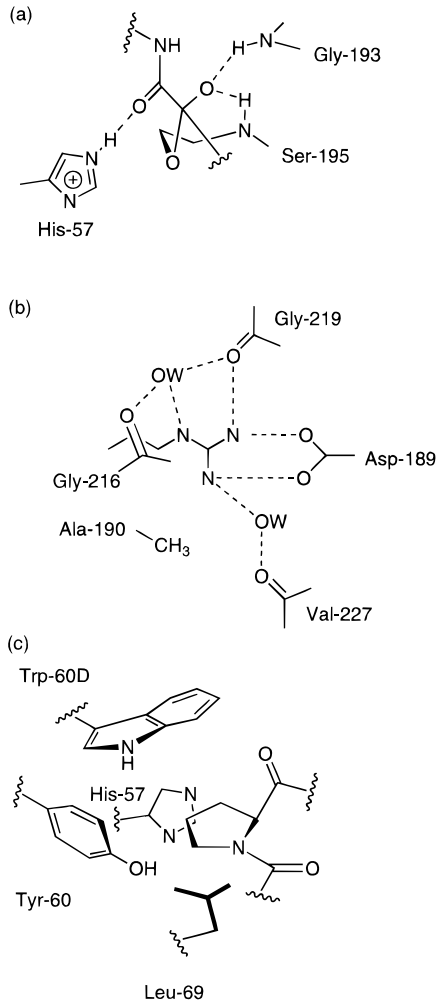


Figure SP*10. (a) Interactions of α-keto amide of **CtA** with catalytic triad include the hemiketal oxygen interacting with the oxyanion hole and the amide carbonyl forming a hydrogen bond with NE2 of His-57. (b) Interaction of guanidine side chain of **CtA** with thrombin S_1 specificity pocket. (c) Interaction of the proline ring of **CtA** with the S_2 subsite of thrombin.

teractions within a rigid macrocyclic framework. This complementarity of the enzyme−inhibitor complex leads to a deep energy minimum for the complex. This situation is very similar to the complexes formed by eglin C and related protein based inhibitors.

Another macrocyclic marine natural product cyclotheonamide A (**CtA**) has been reported to be a potent slow-binding inhibitor of both thrombin (K_i = 1.0 nM) and trypsin (K_i = 0.2 nM).[147,148] The structure of both of these complexes has been reported (trypsin: PDB entry 1TYM; thrombin PDB entry 1TMB). We will first describe the complex with trypsin[147] and then note the similarities and differences in the complex with thrombin.[148,149]

The structure of the cyclotheonamide trypsin complex (see Figure SP*10) shows that the nucleophilic serine-195 hydroxyl group has attacked the ketone carbonyl of the α-keto amide functionality. The hemiketal oxygen (protonation state unknown, and possibly the oxyanion) makes hydrogen-bonding interactions in the oxyanion hole. The amide carbonyl makes a hydrogen bond to His-57, which is presumably protonated. As a result of the carbonyl suffering nucleophilic addition, the hydrogen bond between the

P_1 NH and the backbone carbonyl of Ser-214 is 2.95 Å. For **A90720A**, which does not suffer addition, that hydrogen bond is longer (3.1 Å).

The interaction of the arginine of **CtA** in the S_1 specificity pocket is nearly identical to that observed for **A90720A**. The *N*-formyl diaminopropionamide group of **CtA** forms substrate like hydrogen bonds to Gly-216. The vinylogous tyrosine residue acts as a rigid bridge between the P_3 and P_1' residues and is mostly solvent exposed. The phenol group of this residue is completely solvent exposed. An transannular hydrogen bond between the proline carbonyl and the β-amido group of the diaminopropionamide may serve to rigidify the macrocycle. The P_1' D-phenylalanine side chain makes VDW contacts on the surface of the protein with Tyr-39 and Phe-41; in addition, it makes an edge-to-face with the vinylogous tyrosine side chain. The proline ring makes VDW contacts in the S_2 subsite.

Compound **CtA** also inhibits thrombin; however, it shows a 5-fold decrease in activity. Thrombin, like trypsin, is a member of the chymotrypsin super family of serine proteases. Both of these enzymes show specificity for basic amino acids in the P_1 position of their substrates. While the catalytic triad and oxyanion hole are nearly superimposable for these two enzymes and the S_1 specificity pockets are very similar, there are also major differences in the two enzymes. In the active-site region the S_2 pocket of the two enzymes are quite different. Trypsin has a mostly solvent-exposed groove, while thrombin has a well-defined hydrophobic pocket. The S_1 specificity pockets of the two enzymes are very similar. The major difference is a the substitution of an alanine, in thrombin, for a serine, in trypsin, at residue 190. This results in the loss of one hydrogen bond between the guanidine of the inhibitor and the protein. Thrombin is more selective for substrates with an arginine, rather than a lysine, P_1 residue than is trypsin.[150]

The S_2 subsite of thrombin, unlike trypsin has a well-defined shape. The proline ring binds in hydrophobic pocket formed on three sides by aromatic residues, Tyr-60, Trp-60D, and His-57 and on another side by Leu-69. This pocket appears quite complimentary to the proline ring.

The major difference in the ligand between the two complexes is the conformation of the vinylogous tyrosine unit. For the trypsin complex, the backbone exists in an s-cis conformation, while in thrombin it has an s-trans conformation. The s-trans conformation results in a longer transannular hydrogen bond between the proline carbonyl and the β-amido group of the diaminopropionamide (3.1 Å in thrombin vs 2.8 Å in trypsin). The side chains of this unit and the D-phenylalanine also have different conformations in the two complexes. These different conformations presumably allow each of these side chains to make their optimal VDW contacts with each respective enzyme. The two enzymes are structurally distinct in these regions. It has been proposed that the less than optimal interactions of these two groups with thrombin are partly responsible for the increased specificity of trypsin for **CtA**.[149] However, despite these sub optimal interactions, **CtA** is still a very potent thrombin inhibitor.

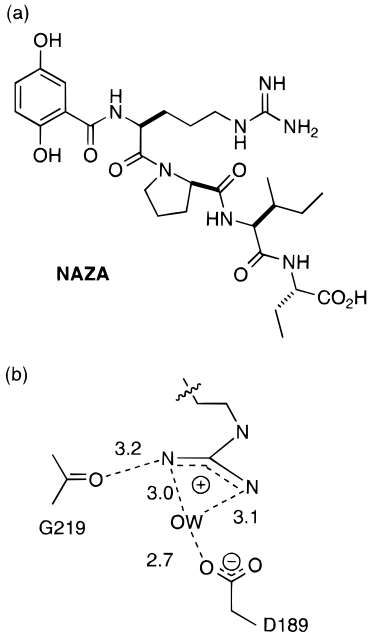


Figure SP*11. (a) Structure of the marine natural product **NAZA** and (b) interaction of guanidine side chain of **NAZA** with the thrombin S_1 specificity pocket.

The structure between thrombin and the natural product Nazumamide A (**NAZA**) has been reported. Natural product **NAZA** has an $IC_{50} = 4.6\ \mu M$ against thrombin but does not inhibit trypsin at concentrations up to 165 μM.[151] This structure showed that **NAZA** is a "retro-binder"; the C-terminal α-aminobutyrate is in the S_4 subsite, the isoleucine in S_3, the proline in the S_2 subsite, and the arginine in S_1. The 2,5-dihydroxybenzene is disordered and located in the S_1' subsite. Except for a long polar interaction (3.7 Å) between the carbonyl oxygen of the proline of **NAZA** and the backbone NH of Glu-216 there are no hydrogen bonds between the backbone of the ligand and the protein. The major forces holding the complex together appear to be interactions of the side chains with their respective subsites.

In the S_1 specificity pocket, a novel hydrogen-bonding arrangement is observed. (See Figure SP*11.) Due to the "retro-binding" nature of **NAZA** the guanidine side chain is unable to make a direct interaction with Asp-189. Thus, unlike many other thrombin inhibitors, it makes no direct hydrogen bond to the carboxylate of Asp-189, but rather a water-mediated interaction.

While the aromatic ring makes interactions with the S_1' subsite, including less than optimal aromatic–aromatic interactions, it does so in two different orientations. It has been proposed that these interactions are responsible for the thrombin selectivity of this natural product.[151]

The proline ring of **NAZA** makes VDW contacts with the S_2 subsite. The isoleucine side chain makes VDW contacts with the hydrophobic portion of the side chain of Glu-213. The C-terminal group is disordered and does not appear to make favorable contacts with the protein. There does not appear to be any protein group to interact with the negative

charge of the carboxylate, but is rather forced into hydrophobic regions of the protein.

The structure of the complex between **NAZA** and thrombin exhibits a novel binding mode. The binding forces appear to be interactions of the side chains of the inhibitor with S_1, S_2, S_3, and S_1' as well as many suboptimal interactions. It was suggested[151] that modifications of **NAZA** which remove these suboptimal interactions could lead to more potent, selective thrombin inhibitors. Thus, the natural product, and its cocrystal structure have provided a new lead for thrombin inhibition.

C. Transition-State Analogs

A common characteristic of many serine protease inhibitors is that they form a covalent bond with the nucleophilic serine hydroxyl of the enzyme. These complexes are often described in the literature as transition-state analogs.[152]

The natural products chymostatin and leupeptin are representative of peptidyl aldehyde inhibitors of serine proteases. This class of compounds are often referred to in the literature as transition-state mimics. It has been reported that the inhibitory potency of two aldehyde based inhibitors of elastase increased by 0.74 times the change in k_{cat}/K_m observed in analogous substrates.[153] However, as was observed in the structure of these complexes the aldehyde does not undergo nucleophilic attack in a single distinct manner. NMR studies of the complex between *N*-acetyl-L-phenylalanal and α-chymotrypsin also show two distinct covalent adducts.[154] The carbonyl carbon of a true transition-state mimic would be expected to undergo nucleophilic attack in a single manner, with the carbonyl oxygen in the oxyanion hole.

The structure of the complex between aldehyde inhibitor **SP-6** and SGPA has also been reported (PDB entry: 3SGA).[155] In this complex the nucleophilic serine forms a covalent adduct with the aldehyde to form a single diastereomeric hemiacetal adduct. The hemiacetal hydroxyl group in this case interacts with the oxyanion hole. In this complex a very large conformational movement of the imidazole ring of His-57 was observed. The imidazole group does not interact with the ligand but rather has rotated out of the catalytic triad and into solvent. The histidine side chain is replaced by a two water molecules, one water hydrogen bonds to Asp-102.

These structures thus suggest that peptidyl aldehydes are not true transition-state analogs, but rather the aldehyde group provides a covalent anchoring mechanism. Peptide aldehyde inhibitors are converted by the enzyme into stable adducts that lack the requisite functionality (i.e., a leaving group) for processing, and are trapped in potential energy wells. These types of inhibitors are referred to as "dead end" inhibitors.[152]

Trifluoromethyl ketone (TFK)-based peptidyl inhibitors have been shown to mimic the high-energy tetrahedral intermediate for amide hydrolysis.[156] Inhibitor **SP-7** is a reversible, competitive, slow-binding inhibitor of chymotrypsin ($K_i = 1.2\ \mu M$). The structure of the complex between **SP-7** and chymotrypsin has been determined by crystallography (PDB entry: 7GCH),[157] and the protonation state of key

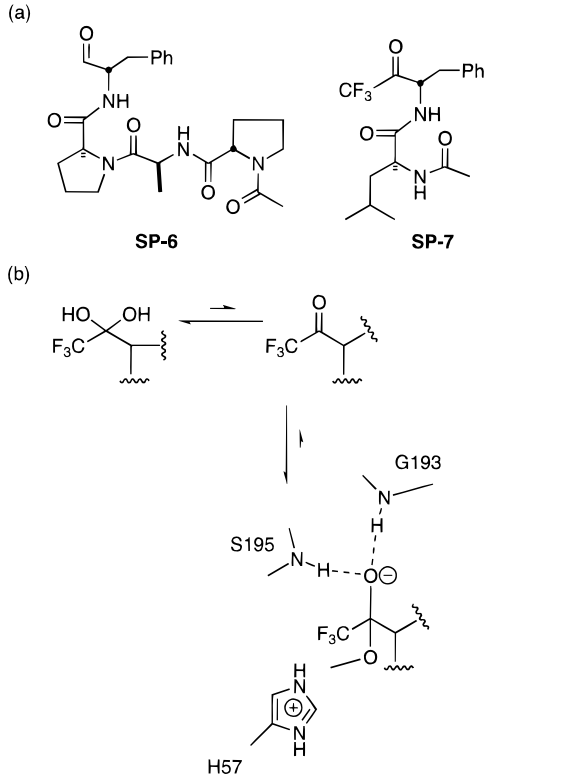


Figure SP*12. (a) Structures of various inhibitors. (b) TFK inhibitors exist in solution mostly as their hydrate. With serine proteases TFK's form covalent hemiketal adducts which exist as the oxyanion and the imidazole ring of His-57 is doubly protonated and charged. Thus, despite an unfavorable solution equilibrium TFK inhibitors are potent inhibitors of serine proteases due to highly favorable electrostatic interactions. In contrast aldehyde inhibitors appear to bind as neutral species.

residues determined by NMR spectroscopy.[158] The structure shows that the nucleophilic serine has attacked the ketone carbonyl of the inhibitor to form a tetrahedral hemiketal. (See Figure SP*12.) The hemiketal oxygen is located in the oxyanion hole and forms hydrogen bonds to the backbone NH's of Ser-195 (2.74 Å) and Gly-193 (2.75 Å). From NMR experiments it was concluded that the inhibitor is bound as an ionized hemiketal. This oxyanion is stabilized formation of two hydrogen bonds to the enzyme. These NMR studies also concluded that the imidazole of His-57 is doubly protonated. This work suggests that TFK analogs resemble the high-energy tetrahedral intermediate in peptide hydrolysis. It was proposed that the resemblance of this adduct to a tetrahedral intermediate is probably a major factor contributing to tight binding. One of the fluorine atoms interacts with Nϵ2 of His-57 at a distance of 3.3 Å, possibly forming a dipole−dipole interaction. NMR studies have also shown that **SP-7** exists in water mainly as the hydrate. This unfavorable solution equilibrium is detrimental to tight enzyme inhibition. When adjusted for TFK concentration the K_i for **SP-7** is 0.53 nM.

The remainder of inhibitor **SP-7** binds in a substrate-like manner. The benzyl group make hydrophobic contacts in the S_1 specificity pocket. The leucine side chain makes hydrophobic contacts with His-57 and Ile-99 in the S_2 subsite, the acetyl carbonyl makes a hydrogen bond to the backbone NH

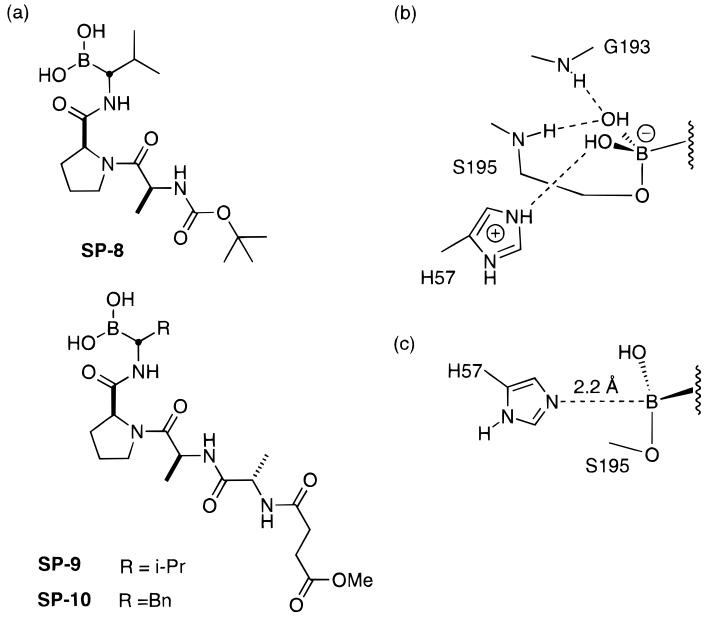


Figure SP*13. (a) Structures of various boronic acid inhibitors. (b) Type I complex of a boronic acid with a serine protease. A charged tetracoordinate boronate species interacts with a charged histidine. (c) Type II complex of a boronic acid with serine protease. A neutral tricoordinate boron species interacts with a neutral histidine.

of Gly-216, and the P_1 amide NH forms a long (3.4 Å) hydrogen bond to the backbone carbonyl of Ser-214.

For a series of seven TFK inhibitors of chymotrypsin, inhibitor properties were compared to kinetic constants of homologous peptide methyl ester and peptide amide substrates.[157] Plots of log inhibitor K_i vs log substrate (k_{cat}/K_m) are linear with slopes of 0.65, indicating transition-state analogy. The slope indicates that these inhibitors are able to utilize 65% of the total binding energy between chymotrypsin and its hydrolytic transition state. The slow binding behavior is proposed to be due to slow formation of VDW contacts between the P_2 side chain and the S_2 subsite.

Peptide analogs containing an α-amino boronic acid in the carboxy terminal position have been shown to be potent reversible inhibitors of serine proteases.[159,160] Compound **SP-8** inhibits α-lytic protease with a $K_i = 0.35$ nM. The structure of the complex between **SP-8** and α-lytic protease (ALP) has been reported (PDB entry: 1P01).[161] The nucleophile Ser-195 has formed a covalent bond to the boron atom forming a tetracoordinate boronate species. The geometry of the adduct is distorted from tetrahedral toward a trigonal planar arrangement with Ser-195 occupying an axial position. The HO−B−OH angle is 115°. Boron angles of 109° would be expected for a tetrahedral adduct, while 120° angles would be expected for a trigonal-planar adduct. Nucleophilic attack on boron produces a boronate in which negative charge resides on boron. One of the oxygens on boron is in the oxyanion hole and makes hydrogen bonds to the backbone NH's of Ser-195 and Gly-193. The other oxygen makes a hydrogen bond to His-57 which is protonated. ^{15}N NMR experiments have shown that His-57 is protonated in the complex.[162] The high potency of peptide boronic acids (a boronic acid binds between 3 and 4 orders of magnitude more tightly than the analogous aldehydes) is likely due to two factors: (1) The ability to simultaneously form hydrogen bonds in both the oxyanion hole and with imidazole of His-57. (2) The favorable electrostatic interaction between the negatively charged boron atom and the protonated imidazole. The remainder of **SP-8** binds in a substrate-like fashion, with the Boc group in the P_4 subsite. Boronic acid complexes with binding characteristics similar to **SP-8** are referred to as type I complexes. (See Figure SP*13b.)

The more extended boronic acid **SP-9** is a very potent ALP inhibitor with a $K_i = 6.4$ nM. Substitution of the P_1 isopropyl side chain with the larger benzyl group results is a decrease in potency ($K_i =$ 540 nM). The structure of each of these complexes have been reported (**SP-9** PDB entry: 1P03; **SP-10** PDB entry: 1P06).[163] The structure of the **SP-9** complex was superimposable with the **SP-8** complex from P_3 to P_1. However, the **SP-10** complex was quite different than the **SP-9** complex. The P_1 benzyl group is too large and inflexible to be incorporated in the S_1 subsite of ALP. As a result, the interaction of the boronic acid with the hydroxyl group of Ser-195 is very different. Rather than a tetracoordinate charged boronate species, a tricoordinate neutral boron species is observed. In this case the boronic acid has undergone monoesterification with the hydroxyl group of Ser-195. The NE2 nitrogen of histidine 57 now coordinates with the boron atom at a distance of 2.2 Å. The geometry at boron is distorted trigonally with an O−B−O angle of 110° and an O−B−C angle of 124.0°. In this structure, a water molecule occupies the oxyanion hole. Boronic acid complexes with binding characteristics similar to **SP-10** are referred to as type II complexes. (See Figure SP*13c.)

When the size of the S_1 subsite was expanded by the replacing Met-213 with alanine, the bound conformation inhibitor **SP-10** changed and now resembled **SP-9**[161,163−166] (PDB entry: 8LPR). This point mutation now resulted in a tetracoordinate boronate species characteristic of type I complexes.

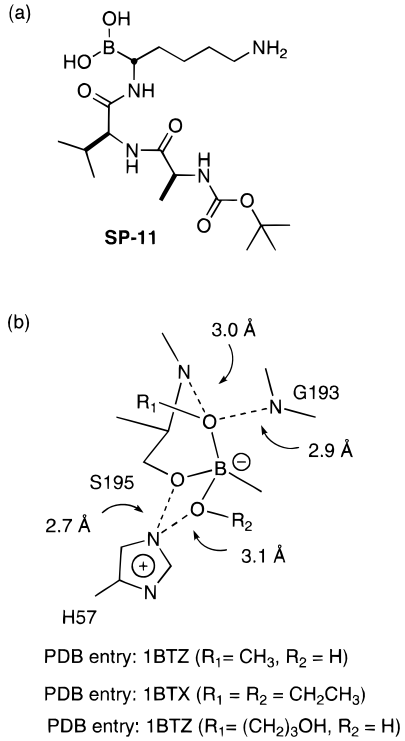


Figure SP*14. (a) Structure of trypsin inhibitor **SP-11**. (b) Crystallization of **SP-11** and trypsin in the presence of different alcohols results in different complexes in which esterification of one or both of the boronate oxygens occurs by reaction with the alcohol from the crystallization medium. This process of assembling boronate esters in the active site has been referred to as episelection.

For a series of 17, out of 19, inhibitors a linear correlation of log inhibitor K_i vs log substrate (k_{cat}/K_m) with a slope of 0.69 was determined, indicating transition state analogy. The slope indicates that these inhibitors are able to utilize 70% of the total binding energy between α-lytic protease and its hydrolytic transition state.[164]

A recent report described the crystallization of the tripeptide boronate **SP-11** with trypsin (K_i = 7 nM) in the presence of 7.5% methanol.[167] The resulting solved structure showed that a stereoselective esterification of one of the boronate oxygens had occurred by reaction with methanol from the crystallization medium (PDB entries: 1BTW, 1BTX, 1BTZ). (See Figure SP*14.) Only the oxygen which is located in the oxyanion hole is esterified, the oxygen which interacts with His-57 is not esterified. Similarly, crystallization of **SP-11** in the presence 1,3-propanediol resulted in a stereoselective esterification of one of the boronate oxygens. In contrast, crystallization of **SP-11** in the presence of ethanol resulted in esterification of both of the boronate oxygens. These esterification processes either occur by reaction in solution followed by the most compatible ester being selectively bound by the enzyme (epitaxial selection), or by being assembled by epitaxial reaction on the enzyme surface. This process of assembling novel boronate esters has been referred to as episelection.

For these boronate ester complexes, the geometry at the boron atom is intermediate between tetrahedral and trigonal. For both the methyl monoester and the ethyl diester the orientation of the boronate group, relative to the catalytic triad, is slightly altered compared to the structure of **SP-8** and α-lytic protease. For both adducts, the oxygen which is oriented in the oxyanion hole has been esterified. These oxygens hydrogen bond to the backbone NH's of both Ser-195 and Gly-193. For the methanol adduct, the methyl group attached to this oxygen makes VDW contacts with the protein in the vicinity of the oxyanion hole and the Cys42−Cys58 disulfide and the free hydroxyl makes a hydrogen bond to His-57 (3.1 Å). That hydroxyl group was replaced with an alkyl group only for the case where the crystallization occurred in ethanol. The oxygen which is derived from the alcohol of Ser-195 forms a very close hydrogen bond (2.7 Å) to His-57. For the diethyl ester, the electron density shows disordered electron density for the ethyl group that is not in the oxyanion hole. For the ethyl group which is located in the oxyanion hole, the electron density suggests either disorder or partial occupancy. This ethyl group makes VDW contacts with the Cys42−Cys58 disulfide.

The structure between **SP-11** and trypsin crystallized in the presence of 1,3-propanediol shows a slightly different structure. In this case, only the oxygen which is located in the oxyanion hole is esterified. The hydroxyalkyl group extends into the S_1' and S_2' subsites, although the hydroxypropane unit is somewhat disordered. The terminal hydroxyl group makes hydrogen bonds to the carbonyl oxygen of Phe-41 in one orientation and to a water molecule in a second orientation. In this structure the oxygen in the oxyanion hole is closer to the backbone NH of Gly-193 (2.9 Å) than it is to the backbone NH of Ser-195 (3.9 Å). The hydroxyl group on boron now forms a closer hydrogen bond to His-57 (2.7 Å). Overall weaker electron density was observed for this inhibitor. This suggests more disorder involving the tripeptide, which nevertheless samples the environment from S_4 to S_2'.

For boronic acids in alcohol−water solutions an equilibrium exits between the free acid and mono- and diesters of the boronate. In these fascinating examples, the enzyme selects the component(s) with the highest affinities out of the equilibrium mixture. This "epitaxial selection" process on the enzyme surface provides insight into previously unrecognized binding sites. Since alcohols and boronic acids are in equilibrium boronic esters, it is also significant that the covalent interaction between the Ser-195 hydroxyl and the more potent boronic acids involves formation of the tetracoordinate boronate species as opposed to a tricoordinate neutral boron monoester. This information suggests that the charged tetracoordinate boronate species interacting with a charged histidine is more stable than a neutral tricoordinate boron species interacting with a neutral histidine.

Structures between serine proteases and peptidyl aldehyde, trifluoromethyl ketones and boronic acid inhibitors only provide information regarding how residues to the amino terminal side of the scissile bond interact with the enzyme during the "transition state". Structural information regarding the binding of residues to the carboxy terminal side of the scissile bond (i.e., the "prime side") is not available from these

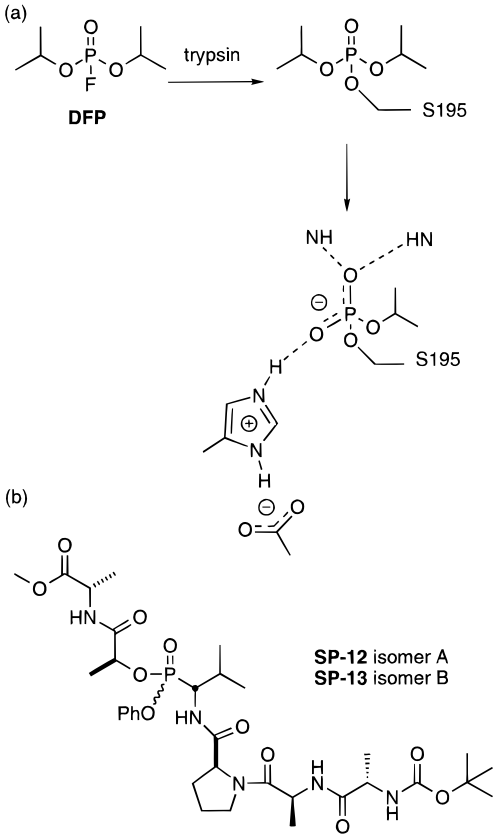


Figure SP*15. (a) Compound **DFP** irreversibly reacts with the active-site serine of serine proteases to give a neutral diisopropyl phosphoryl group. Slow hydrolysis of one of the isopropoxy groups results in a charged phosphate. (b) Peptidic analogs of **DFP** as transition-state mimics of serine proteases. These compounds are isomeric at the phosphorous stereocenter.

classes of compounds. Structures of complexes between serine proteases and protein-based inhibitors (i.e., eglin C) reveal how substrates interact with these enzymes on both sides of the scissile bond. However, these compounds to not form a covalent bond to the enzyme and can only be considered models for a Michaelis complex, not a tetrahedral intermediate or a transition state analog. Structural studies with peptidic phosphonylating agents, which act as irreversible inhibitors of serine proteases, provide structural information about residues on the "prime side" of a covalent adduct.

Diisopropyl phosphofluoridate (**DFP**) is a potent, nonselective, serine protease inhibitor which irreversibly reacts with the active-site serine hydroxyl group.[168−171] Crystallization of **DFP** with trypsin results in the active-site serine being covalently bound to a monoisopropyl phosphoryl group. This product results from initial phosphonylation of the active site serine to give a neutral diisopropyl phosphoryl group, followed by hydrolysis of one of the isopropoxy groups to give a charged phosphate. (See Figure SP*15.) The structure of the final complex was solved by neutron diffraction (PDB entry: 1NTP).[172,173] The unique ability of neutron diffraction to experimentally locate the position of hydrogen atoms allowed the determination of the protonation state of key active-site residues. In this complex, the histidine (H57) is charged and doubly protonated, and the aspartate (D102) is charged and unprotonated. In the complex, the charged phosphate forms a hydrogen bond to NE2 of His-57 with one oxygen and with the other oxygen it makes hydrogen bonds in the oxyanion hole with the backbone NH's of Gly-193 and Ser-195. The isopropyl group is at the opening of the S_1 subsite.

Peptidic analogs **SP-12** and **SP-13** (stereoisomers at phosphorous) of **DFP** have been reported to be irreversible inhibitors of α-lytic protease.[174] One isomer (**SP-12**) shows a factor of 2 increase in the second-order rate constant for enzyme inactivation. The structure of each complex with α-lytic protease has been reported (PDB entries: 1P11 and 1P12).[175]

The complex between **SP-13** and α-lytic protease shows, as expected, a covalent bond between the hydroxyl group of Ser-195 and the phosphorus atom. This results in a tetrahedral adduct. The phosphoryl oxygen is in the oxyanion hole and makes hydrogen bonds with the backbone NH's of Ser-195 and Gly-193. The inhibitor spans both sides of the scissile bond and interacts with the seven substrate recognition sites P_5 to P_2'). Two hydrogen bonds between the P_2' carbonyl and backbone NH and the backbone NH and carbonyl of Leu-41 are observed. The methyl side chains of the P_1' and P_2' groups make solvent-exposed VDW contacts with the protein. The conformation and ligand protein interactions of the P_1' and P_2' residues are similar to those seen in protein based inhibitors. No protein conformation changes are observed to accommodate the "primed" residues. The major unexpected feature in this complex is the conformation of the side chain of His-57. The imidazole group does not interact with the phosphonate ligand but rather has rotated out of the catalytic triad and into solvent. The histidine side chain is replaced by a single water molecule which hydrogen bonds to Asp-102. This conformational change is similar to that observed for the complex between SGPA and the peptide aldehyde **SP-6**.[155] It was proposed that a positively charged histidine can only be accommodated in a native-like position if there is a negative charge in or near the oxyanion hole.[175] A neutral histidine would have an unfavorable electrostatic interaction with the phosphonate oxygen.

Isomer **SP-12** forms a mixture of two different tetrahedral adducts in the active site, both covalently bound to Ser-195. One adduct, at 58% occupancy, is exactly the same as the complex with **SP-13**. The other adduct, at 42% occupancy, has lost the P_1' and P_2' residues by hydrolysis. This is similar to the case of **DFP**. In the lower occupancy structure, His-57 does not rotate out of the active site, but rather it interacts with the charged phosphonate group.

Compounds **SP-12** and **SP-13** are diastereomers, of unknown configuration at phosphorous, and both lead to the same adduct in the enzyme−ligand complex. Therefore, one isomer must react with Ser-195 with inversion of configuration at phosphorous while the other isomer reacts with the protein with retention of configuration at phosphorous. These structures are very interesting since they serve as a model for the conformation of tetrahedral adduct of a peptide spanning P_5 to P_2'. In addition, the conformation change, observed for His-57 in the neutral inhibitor adduct, suggests that a strong

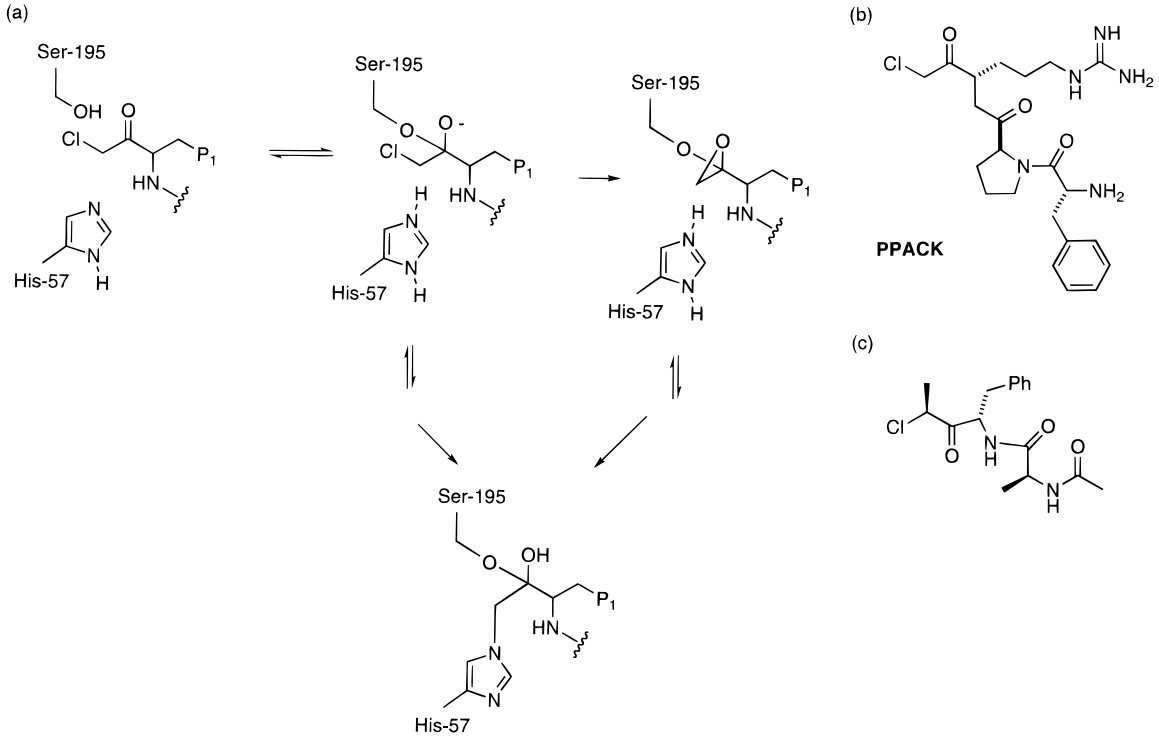


Figure SP*16. (a) α-Chloro ketones alkylate NE2 of His-57 in a process that requires the hydroxyl group of Ser-195. Two mechanisms for alkylation are possible, either direct nucleophilic attack on the α-chloro adduct or nucleophilic attack on an intermediate epoxide. (b) Structure of **PPACK**, a thrombin affinity label. (c) *N*-Acetyl-L-alanyl-L-phenylalanyl-α-chloroethane alkylates chymotrypsin with net retention of configuration.

hydrogen bond, or a favorable charge−charge interaction is needed for His-57 to interact with an inhibitor. With the exception of the histidine conformational change, the complex between **SP-13** and α-lytic protease provides a model for the structure of the transition state (or a high-energy reaction intermediate) that spans both sides of the scissile bond.

D. Thrombin

1. Introduction

Thrombin is a serine protease responsible for the cleavage of soluble fibrinogen into insoluble fibrin in the last protease-mediated step of the coagulation cascade. Compounds which inhibit thrombin are effective inhibitors of blood clotting. As a result, there has been much research aimed at the development of small molecule inhibitors of thrombin as clinically useful anticoagulants.[176]

Huridin, a small protein of 65 amino acids, isolated from the leech *Hirudo medicinalis*, is the most potent natural inhibitor of thrombin known.[177] It inhibits thrombin with a K_i = 22 pM. The key features of thrombin inhibition by hirudin are that the N-terminal binds to the active site in a nonsubstrate mode, the peptide chain runs parallel to the active site residues Ser-214 to Gly-216, and a substantial portion of the C-terminal region (Asp-55 to Gln-65) interacts with thrombin in what is referred to as the fibrinogen recognition exosite[178−180] (PDB entries: 1HGT, 4HTC). A discussion of the protein−protein interactions in the exosite is beyond the scope of this review.

2. Active Site-Based Inhibitors

The α-chloro methyl ketone **PPACK** is a potent irreversible inhibitor of thrombin. The crystal structure of its complex has been reported (PDB entry: 1PPB).[181] The ketone carbonyl has undergone nucleophilic attack by serine-195 to form a hemiketal, and NE2 of histidine-57 has formed a covalent bond by displacing the chlorine atom. It is this bond formation which makes the complex formation irreversible. Unlike the hemiketal oxygen of **CtA**, the hemiketal oxygen of **PPACK** does not make intimate contact with the oxyanion hole. Rather it points in the direction of the oxyanion hole (O−N S195 distance: 3.2 Å; O−N G193 distance: 3.2 Å) but it is somewhat solvent exposed. Two mechanisms for histidine alkylation are possible. (See Figure SP*16.) In the first the histidine nitrogen directly alkylates the chlorine after hemiketal formation and proton transfers. The second possible mechanism involves reversible hemiketal formation, followed by conversion of the adduct into an epoxide, the histidine then alkylates the epoxide after the appropriate proton transfers. The **PPACK** thrombin complex is typical of complexes between serine proteases and α-chloro ketones.

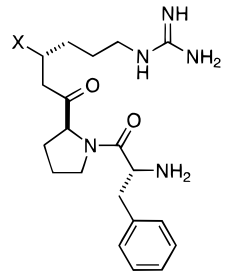
The Brandeis group[182] has reported studies which address this mechanistic issue. This study involved irreversibly inhibiting γ-chymotrypsin with enantiomerically pure (2*S*)-*N*-acetyl-L-alanyl-L-phenylalanyl-α-chloroethane and determining the structure of the inhibited complex by X-ray crystallography. It was determined that histidine alkylation occurred with net retention of configuration. This result is consistent with the presence of an intermediate epoxy ether adduct. The epoxide mechanism is also consistent

with kinetic evidence.[183,184] The stereochemistry of alkylation is inconsistent with direct S_N2 displacement of the chloro leaving group. In addition, previous studies had shown that the nucleophilic serine hydroxyl was necessary for inactivation of chymotrypsin by α-chloro ketone inhibitors,[185] this suggested that the alkylating agent was a species which first reacts with the nucleophilic serine.

The structure of the inhibited complex was, however, very different than expected (PDB entries: 2GMT and 1DWE) on the basis of complexes between other serine proteases and α-chloro ketones. No hemiketal linkage between Ser-195 and the ketone was seen. The P_1 carbonyl is pointing in a direction almost diametrically opposed to the oxyanion hole. The P_1 phenyl group makes less than optimal interactions in the S_1 subsite and the remainder of the inhibitor is mostly solvent exposed. The authors attribute this unusual conformation to steric hindrance between the extra methyl group and the rest of the inhibitor chain. On the basis of the observed net retention in the alkylation process it is likely that inhibition occurs via a process of hemiketal formation, conversion to the epoxide, histidine alkylation of the epoxide and then a conformation change into the observed complex.

For **PPACK**, the P_1 arginine substituent binds in the S_1 specificity pocket in the same manner as the arginine of **CtA**. Like **CtA**, the P_2 proline ring occupies the well-defined S_2 subsite. The P_3 group is D-phenylalanine. Its phenyl side chain makes an edge-to-face interaction with Trp-215 and VDW contacts with Leu-99 and Ile-174. This would normally be the S_4 subsite for a substrate. The D configuration allows a P_3 group to fill the S_4 subsite and in addition allows the N-terminal ammonium group to make a hydrogen bond to the carbonyl of Gly-216. The presence of a D amino acid in the P_3 position has also been observed in the marine natural product **A90720A**. The D-Phe carbonyl makes a long (3.1 Å) hydrogen bond with the backbone NH of Gly-216. In the bound complex the side chains of the phenylalanine and proline stack against each other. NMR evidence indicates that inhibitors that contain the D-Phe-Pro residues have a solution conformation that is similar to its bound conformation.[186] These compounds are thus preorganized for binding to thrombin. This is an example of productive hydrophobic collapse.

The peptide aldehyde **SP-14**, related to **PPACK** has been reported to be a potent (K_d = 1.8 nM), reversible inhibitor of thrombin.[187] The structure of its complex shows, as expected, covalent bond formation with Ser-195 to form a hemiacetal and noncovalent interactions similar to the **PPACK** complex. Compound **SP-15**, lacking the aldehyde group, is a weaker thrombin inhibitor with K_d = 180 nM.[188] The structure of the complex between **SP-15** and thrombin is very similar to the to the complex with **SP-14**.[189] Thus, the covalent interaction between Ser-195 of thrombin and the aldehyde of the inhibitor is worth 2.7 kcal mol^{-1}. Although **SP-15** is less potent than **SP-14**, it still retains significant anticoagulant activity. An important issue in the development of clinically useful thrombin inhibitors is that they be



SP-14 X = CHO K_d = 1.8 nM

SP-15 X = H K_d = 180. nM

Figure SP*17. Structures of various compounds.

selective for thrombin over other serine proteases. It has been suggested that optimal selectivity is achieved when inhibitors derive their binding energy from interactions with unique structural elements of the target enzyme. Compound **SP-15**, unlike **SP-14** (Figure SP*17), derives its binding energy only from noncovalent interactions with thrombin, not from interactions with the catalytic triad common to related serine proteases. For the related enzyme trypsin, **SP-14** has a K_d = 9.1 nM, while **SP-15** is much less potent with K_d = 5.9 μM. SAR studies based upon **SP-15** have shown that avoidance of a covalent interaction between Ser-195 and the inhibitor may be a useful strategy in the design of selective thrombin inhibitors.[189]

The DuPont Merck group[190] has reported the potent boronic acid-derived thrombin inhibitor **DuP714** (K_i = 40 pM). The crystal structure of the complex between **DuP714** and thrombin shows covalent bond formation between the boron atom of the inhibitor and the hydroxyl oxygen of Ser-195 to give a type I complex. This was as expected on the basis of many other complexes between boronic acids and serine proteases. While the boron atom becomes tetracoordinate in the complex, its geometry is more trigonal than tetrahedral. The Ac-(D)Phe-Pro portion of the inhibitor binds in a manner very similar to the thrombin complex with **PPACK**.

The most significant contribution of this study was the preparation, enzyme evaluation, and crystallographic studies of **DuP714** (arginine P_1 side chain) and its amidine, lysine, ornithine, and homolysine P_1 analogs. (See Figures SP*18 and SP*19.) This study provides a detailed structural study on the electrostatics of protein interactions in the S_1 specificity pocket of thrombin (PDB entries: 1LHC, 1LHD, 1LHE, 1LHF, 1LHG).[190]

The guanidino group of **DuP714** forms a bidentate hydrogen-bonding interaction with the carboxylate group of Asp-189. This interaction occurs at the bottom of the S_1 specificity pocket. Each of the terminal nitrogens is nearly equidistant from the syn face of a carboxylate oxygen. In addition, each nitrogen makes an additional hydrogen bond to the protein, one to the backbone carbonyl of Gly-219 and the second, through a buried water molecule, to the carbonyl of Phe-227. The Nε nitrogen forms a water-mediated hydrogen bond to the backbone carbonyl of Gly-219.

The amidino analog of **DuP714** is a less potent thrombin inhibitor by 1.2 kcal mol^{-1}. This compound

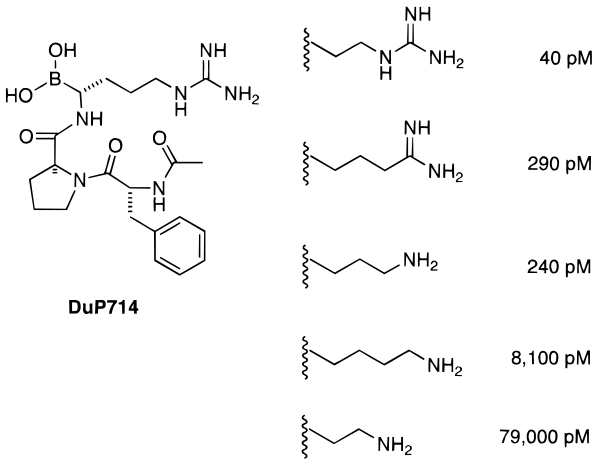


Figure SP*18. Structure of potent thrombin inhibitor **DuP714** and less potent analogs with different P_1 side chains.

binds similarly; however, the substitution of Nϵ with a methylene results in the loss of the water-mediated hydrogen bond. The increased basicity of a guanidine group compared to an amidine group may also contribute to a stronger hydrogen bond to the carboxylate. The carbon for nitrogen substitution may also make the amidine side chain slightly more flexible in solution. Increased ligand flexibility in solution would be detrimental to binding from an entropic point of view.

The substitution of the arginine side chain of **DuP714** with a lysine side chain results in less potent inhibition by 1.1 kcal mol^{-1}. The ammonium group of this ligand does not interact directly with Asp-189. This interaction is mediated by two water molecules. The ammonium group forms direct hydrogen bonds to the backbone carbonyls of both Ala-190 and Gly-219. Unlike the arginine and guanidine side chains the lysine side chain traps two water molecules in the S_1 subsite. The trapping of an additional water molecule is, of course, entropically unfavorable. The equal potency of the lysine and amidine analogs suggest that the favorable interaction made by the lysine side chain with these water molecules and its direct contacts with the protein compensate enthalpically for the cost of trapping an additional water molecule.

The interaction of a lysine side chain with the S_1 subsite in trypsin is somewhat different. The substitution of Ala-190 in thrombin for a serine in trypsin results in subtle differences in the interaction of the lysine side chain with these two proteins. The ammonium group hydrogen bonds to the side chain of Ser-190 in trypsin, not the main chain carbonyl, as observed in thrombin. This results in the side chain being more deeply buried in the protein interior. A water-mediated hydrogen bond to Asp-189 is also observed in trypsin; however, this involves the opposite aspartate oxygen than does thrombin.

Substitution of the arginine with a homolysine side chain resulted in a substantial loss in affinity (3.2 kcal mol^{-1}) for thrombin. The ammonium group makes direct hydrogen bonds to Asp-189 with the nitrogen about 3 Å from each carboxylate oxygen. This nitrogen also forms longer hydrogen bonds to the backbone carbonyls of Ala-190 and Gly-219. A weak, water-mediated hydrogen bond is made to the backbone carbonyl of Phe-227. While the homolysine analog makes a direct interaction with the aspartate and traps only one water molecule, it has 33-fold less affinity for thrombin than the lysine analog (2.1 kcal mol^{-1}). It has been suggested that the entropically unfavorable immobilization of an additional carbon−carbon single bond is largely responsible for the decreased affinity. In addition, it is possible that the water-mediated interactions observed between the lysine and aspartate are highly favorable.[191]

The substitution of the arginine with an ornithine side chain resulted in a substantial (4.5 kcal mol^{-1}) loss in affinity for thrombin. In this case, the shorter side chain results in a less than optimal placement of the ammonium group relative to the carboxylate group (4.3 and 4.5 Å N−O distances). This results in longer water-mediated hydrogen bonds to Asp-189. It is likely that the poorer affinity of the ornithine analog is due to the increased distance of the ammonium ion from the carboxylate and the resulting longer water mediated hydrogen bonds.

This study by the DuPont Merck group and their insightful discussions is very enlightening. This study demonstrates that less than optimal protein−ligand charged interactions can result in quite reduced affinity of the ligand for the protein. It is also clear from this study that obtaining optimal electrostatic interactions between two charged groups can be due to subtle structural features.

The DuPont Merck group has reported studies that use the structures of the **DuP714** thrombin complex, and related structures, to design constrained boropeptide inhibitors (Figure SP*20). In an effort to constrain the conformation of the aromatic ring of **DuP714** analogs and to maximize its interactions with the S_4 subsite of thrombin, constrained analogs related to **SP-16** were designed. Compound **SP-16** (the boronic ester is converted to the boronic acid under the assay conditions) was found to be a potent thrombin inhibitor with K_i = 320 pM.[192]

Another study reported **SP-17**, containing a conformationally restricted benzoic acid subunit, which has K_i = 70 pM against thrombin.[193] The data for related analogs suggested that increased potency of these compounds was not due to preorganization of the P_2 and P_3 groups. It was suggested that inhibitors related to **SP-17** adopt a unique binding mode which allows for the direct interaction of the P_1 lysine side chain with the Asp-189 side chain in the S_1 specificity pocket. No crystal structure of these complexes has been reported.

The DuPont Merck group has also reported that the nature of the N-terminal groups in these inhibitors has a significant effect on their potency as thrombin inhibitors. Compound **DuP714** terminating in an acetyl group had a K_i = 41 pM, the analog terminating with the free amine was more potent with K_i < 3.6 pM and the analog terminating with a Boc group was also more potent with a K_i = 3.6 pM.[194] A series of the corresponding peptide methyl esters were prepared and the same trend of relative potency based on the N-terminal group was observed (Ac-D-

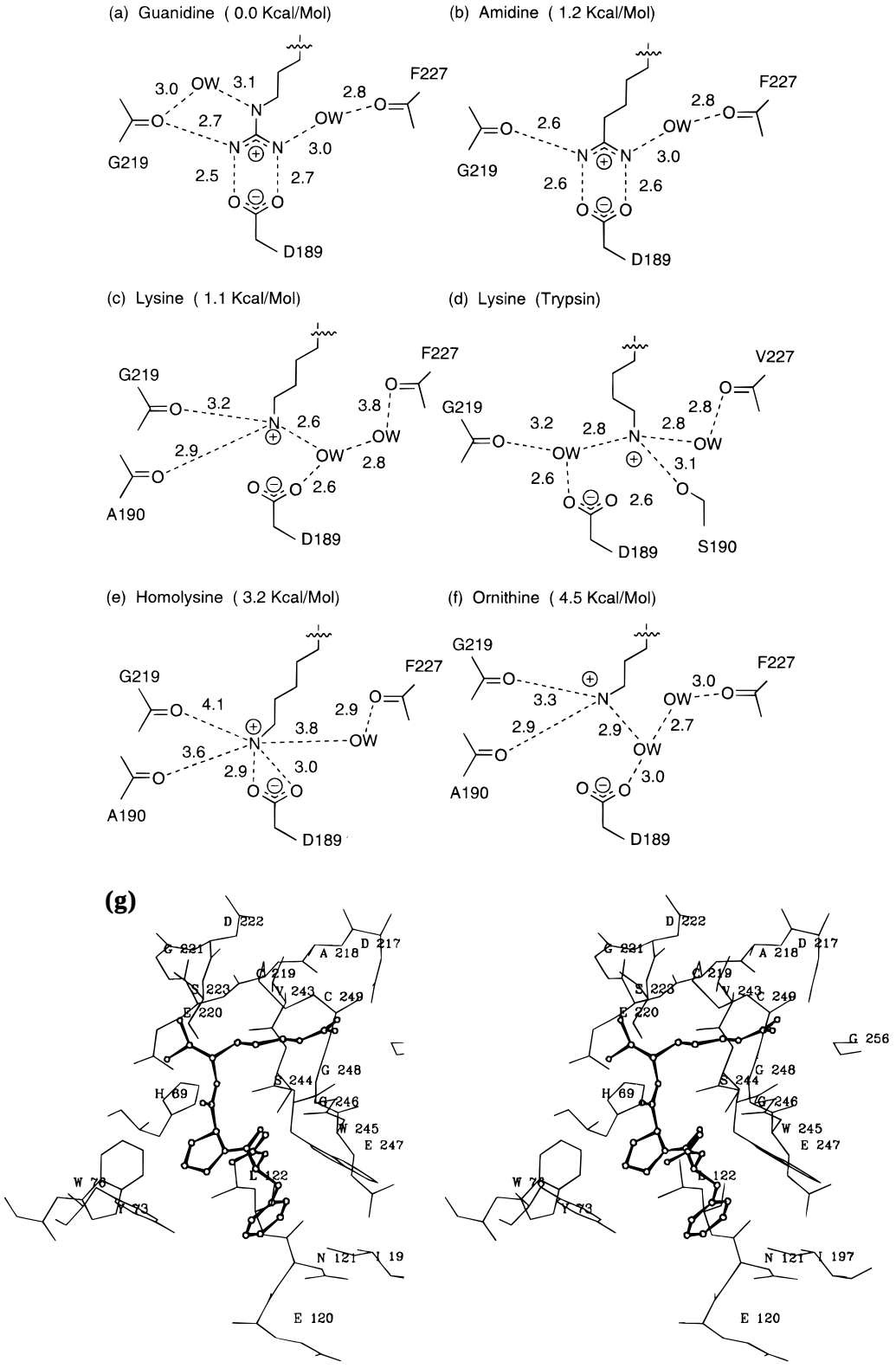


Figure SP*19. Interactions of P_1 side chains of **DuP714** and analogs with S_1 specificity pocket of thrombin: (a) guanidine side chain, (b) amidine side chain, (c) lysine side chain, (d) interaction of a lysine side chain with the related enzyme trypsin, (e) homolysine side chain, and (f) stereoview of **DuP714** thrombin complex.

Phe-Pro-Arg-OMe, $K_i = 60\ \mu$M; Boc-D-Phe-Pro-Arg-OMe, $K_i = 7.8\ \mu$M; H-D-Phe-Pro-Arg-OMe, $K_i = 0.58\ \mu$M). These esters were incubated with thrombin and the structures of resulting complexes, the enzyme-bound product carboxylates, were determined by crystallography.[195] Examination of these crystal structures suggested that the higher affinity of the unblocked peptide is due to a different conformation of side chain of Glu192 in order to maximize its electrostatic interaction with the terminal ammonium group of the inhibitor. The distance between the ammonium group and the carboxylate is however 4.4 Å. At this distance, this now becomes a solvent-mediated salt bridge. The relatively long distance between these two ions reminds one that electrostatic interactions are still important over long distances.

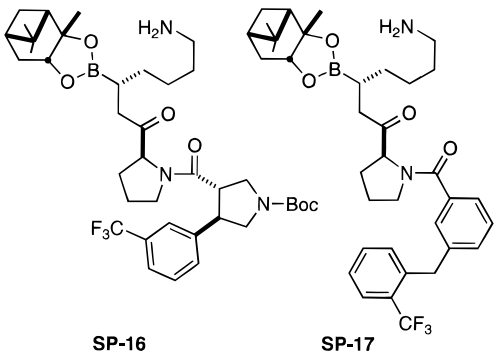


Figure SP*20. Structures of constrained analogs of **DuP714**.

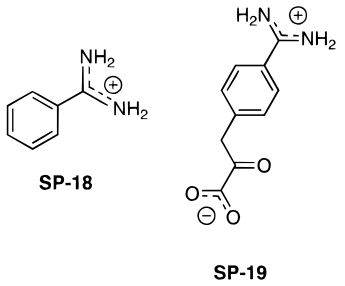


Figure SP*21. Structures of various inhibitors.

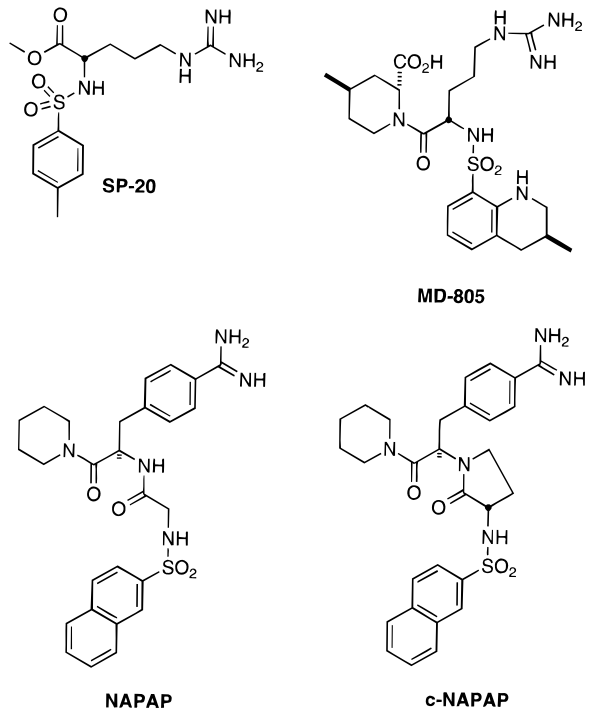


Figure SP*22. Structures of various inhibitors.

In classical electrostatics, charge−charge interactions vary as a function of $1/r$, while noncharged polar interactions vary as a function of $1/r^3$.[196] The greater potency of the Boc vs acetyl groups was due to more extensive VDW contacts with the larger Boc group with the protein, especially with the hydrophobic portion of the side chain of Glu-217. Also observed in the complex with the Boc inhibitor was an altered interaction of the carboxylate of the product inhibitor with the oxyanion hole. This result indicated that changing the structure of the inhibitor can result in altered binding at an unaltered distal site.

Structures of the simple inhibitors of thrombin, benzamidine (**SP-18**) (K_i = 300 nM; PDB entry: 1 DWB)[197] and *p*-amidinophenylpyruvate (**SP-19**) (K_i = 620 nM; PDB entry: 1AHT),[198] have been reported (Figure SP*21). In both of these complexes the benzamidine ring makes VDW and electrostatic contact with the S_1 specificity pocket. For **SP-19**, the ketone carbonyl has undergone nucleophilic attack by Ser-195 with the carbonyl oxygen in the oxyanion hole. The carboxylate group forms a hydrogen bond to His-57.

In a medicinal chemistry program aimed at developing small molecule thrombin inhibitors, the methyl ester of Nα-tosylated arginine, **SP-20**, was used as a starting point for inhibitor elaboration.[150] Compound **SP-20** is an inhibitor that is effectively a poor substrate. These studies ultimately identified **MD-805** (argatroban; K_i = 19 nM)[199] and **NAPAP** (K_i = 6.6 nM)[200] as potent reversible inhibitors of human thrombin (Figure SP*22). Crystal structures for the complex between both **MD-805** and **NAPAP** with human thrombin have been determined by two groups[181,197,201−203] (PDB entries: 1DWC and 1DWD).

Modeling studies[204] proposed a "substrate-like" binding mode for **MD-805**. (See Figure SP*23.) This involved placing the arginine side chain in the S_1 subsite and its carbonyl group in the oxyanion hole with the flanking groups in the S_2 and S_1' subsites. The observed crystal structure showed an unexpected binding mode. (See Figure SP*23.) The arginine side chain occupied the S_1 subsite, but did so in a manner much different than does **PPACK**. The Nϵ nitrogen makes a hydrogen bond to the backbone carbonyl of Gly-219, one of the terminal nitrogens of the guanidine interacts with both carboxylate oxygens of Asp-189, while the other makes a hydrogen bond to a water molecule. This water is nearly tetrahedrally coordinated and makes a hydrogen bond to the carboxylate group of the inhibitor, and longer hydrogen bonds with the hydroxyl group of the nucleophilic Ser-195 and the backbone carbonyl of Ser-214.

In the **MD-805** thrombin complex, the substituted pipecolinic acid ring and the phenyl ring are stacked upon each other. These stacked rings make VDW contacts in the S_2 subsite. The saturated ring of the tetrahydroisoquinoline ring makes VDW in the S_4 subsite. Compound **MD-805** makes very different polar interactions with the enzyme than does a substrate-like inhibitor. The sulfonamide oxygens are solvent exposed; the amide carbonyl is buried and is 3.0 Å from the backbone NH of Gly-216; one of the carboxylate oxygens forms a hydrogen bond to the hydroxyl of Ser-195 and to a buried water molecule; the other carboxylate is solvent exposed. The carbonyl oxygen of Gly-216 forms hydrogen bonds to the sulfonamide NH (2.85 Å) and the tetrahydroisoquinoline NH (3.0 Å).

The observed stacking of the two hydrophobic ligand groups represents an example of productive ligand hydrophobic collapse.[125] The stacking of these two groups generates a surface which is complimentary to the S_2 and S_4 subsites of thrombin. A small molecule crystal structure of **MD-805**, crystallized from methanol (CSD entry: JEBNUZ), also shows a similar stacking of the two rings.[205] It is possible that

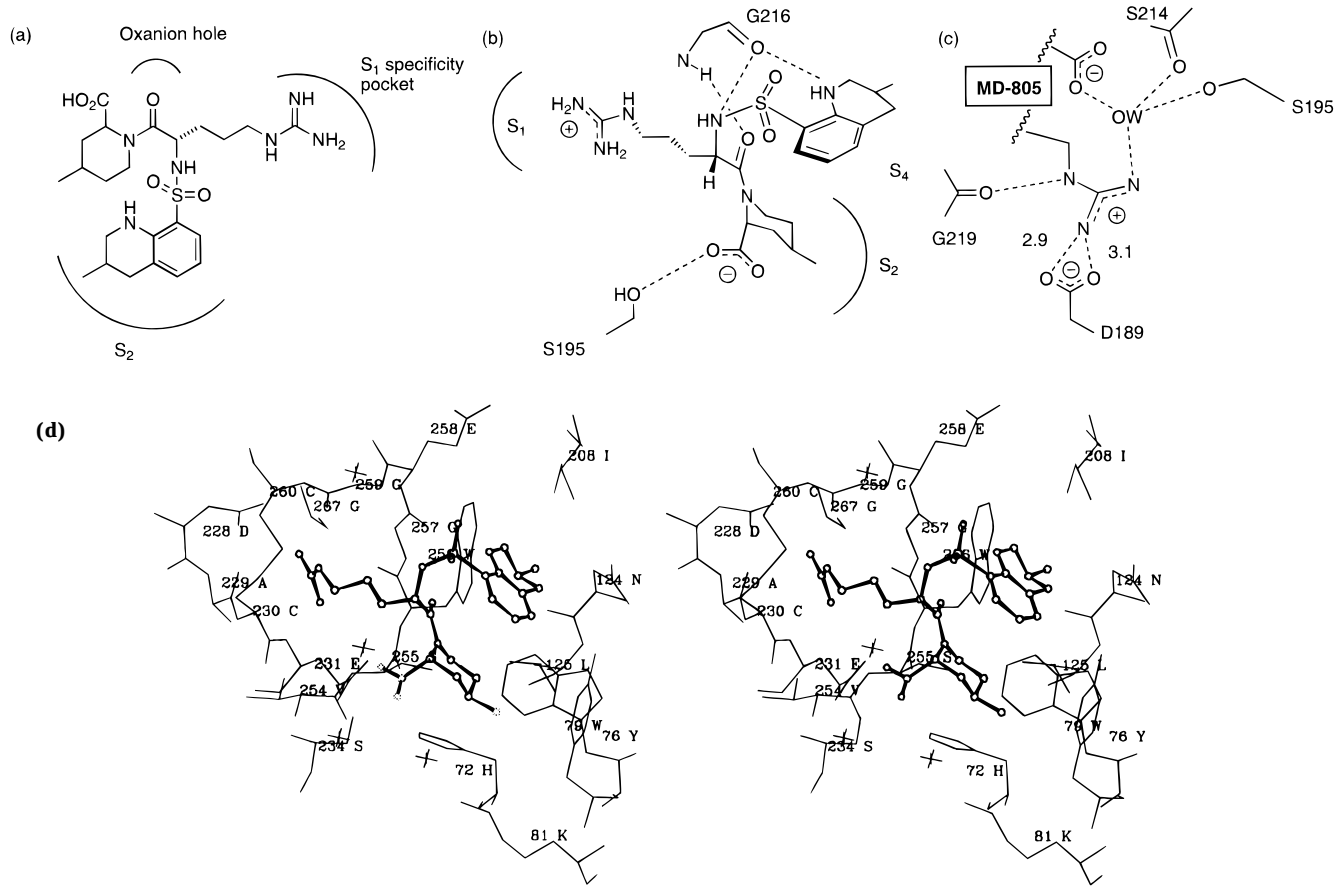


Figure SP*23. (a) Proposed "substrate-like" binding mode for **MD-805**. (b) Observed binding mode for **MD-805**, the two ring systems are stacked upon each other. This appears to be an example of productive hydrophobic collapse. (c) Interactions of **MD-805** in the thrombin S_1 specificity pocket. (d) Stereoview of **MD-805** thrombin complex.

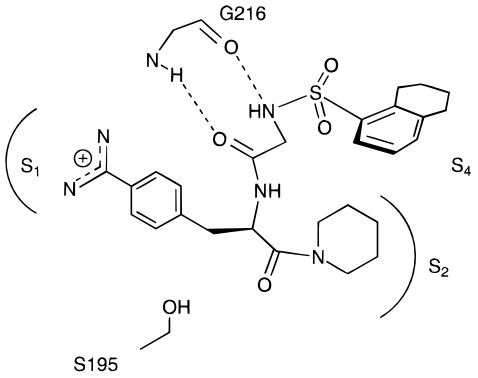


Figure SP*24. Observed binding mode of **NAPAP**.

these groups also interact with each other in aqueous solution and induce conformation which is complimentary to the enzyme.

The inhibitor **NAPAP** binds differently since there is a change in stereochemistry at the center adjacent to the P_1 group. (See Figure SP*24.) The benzamidine group interacts in the S_1 specificity pocket much like benzamidine itself. The piperidine ring makes VDW contacts with the S_2 subsite; however, its carbonyl group is now solvent exposed. The glycine unit of the inhibitor now forms hydrogen bonds to Gly-216 of the enzyme. The sulfonamide oxygens are solvent exposed. The naphthalene ring makes VDW contacts in the S_4 subsite as well as an edge-to-face interaction with Trp-215. For **NAPAP** no interaction is made between the inhibitor and the catalytic triad. As in the case of **MD-805**, ligand hydrophobic collapse is observed. In this example, the naphthalene ring and the piperidine ring stack upon each other.

Both of these compounds are potent thrombin inhibitors without making direct interactions with the catalytic triad. For both complexes, the oxyanion hole is occupied by a water molecule. The nature of the thrombin active site provides for strong electrostatic interactions with the guanidine or amidine group with the S_1 specificity pocket, significant VDW contacts in the S_2 and S_4 subsites as well as additional ligand protein hydrogen bonds. The well-defined shapes of the S_2 and S_4 subsites explain the high specificity of thrombin substrates. Compounds **MD-805** and **NAPAP** are much weaker inhibitors of trypsin with K_i's of 5 μM and 690 nM, respectively. The structure of trypsin shows that it does not have well-defined S_2 and S_4 subsites.

The structure of the **NAPAP** complex with trypsin has also been reported[202] (PDB entry: 1PPC). Trypsin, unlike thrombin, does not have a well-defined S_2 subsite. The structure of the **NAPAP**–trypsin complex shows a bound ligand conformation almost identical to that of the thrombin complex. The weaker potency vs trypsin is due to significantly less VDW contacts in the S_2 subsite. In trypsin, it is a shallow solvent-exposed cleft, while in thrombin it is a well-defined hydrophobic binding site. The hydrophobic collapse observed between the aromatic ring and the piperidine ring is also observed in the trypsin complex. This result is consistent with the ligand having undergone hydrophobic collapse in

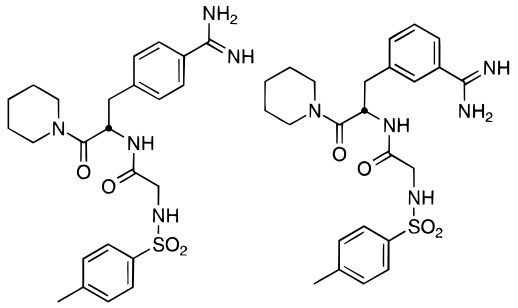


Figure SP*25. Structures of related inhibitors whose structures have been deposited in the PDB.

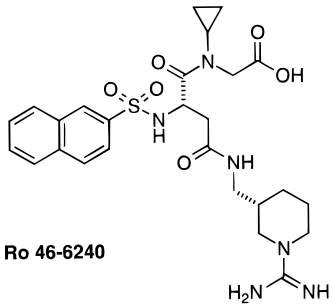


Figure SP*26. Structures of potent thrombin inhibitor **Ro 46-6240**.

solution to produce a bioactive conformation which is then selected by the protein.

Structure of the complexes of related inhibitors (PDB entry: 1ETT and 1PPH) have been reported.[203,206] (See Figure SP*25.)

Using the structure of the complex between **NAPAP** and thrombin as a starting point the BASF group designed and synthesized constrained analogs of **NAPAP**.[207] The complex structure suggested that there was sufficient space for cyclization between the α carbon of the glycine and the nitrogen of the substituted phenylalanine of the inhibitor **NAPAP**. It was expected that this cyclization would lead to higher affinity ligands because these constrained analogs would suffer less loss of configurational entropy upon complex formation. The lactam **c-NAPAP** was prepared and found to be 112-fold more potent as a thrombin inhibitor ($IC_{50} = 1.6$ nM). The structure of the complex between **c-NAPAP** and thrombin showed that the constrained inhibitor **c-NAPAP** bound to thrombin in a mode similar to **NAPAP**.

Further SAR studies, based on the **MD-805** structure, have identified **Ro 46-6240** as a potent and selective inhibitor of thrombin (Figure SP*26). It has a $K_i = 0.71$ nM vs thrombin and 3.80 μM against trypsin.

From a random screening process, compound **SP-21** was identified as an inhibitor of thrombin with a $K_i = 7.5$ μM.[208] Medicinal chemistry optimization of **SP-21** led to **BMS-183507** ($K_i = 17.2$ nM). The structure of its complex with thrombin showed a "retrobinding" fashion for this inhibitor (PDB entry: 1HDT).[209] The inhibitor binds with its alkylguanidine moiety in the S_1 specificity pocket and its two phenyl rings occupying the hydrophobic S_2 and S_4 subsites of the active site. In this arrangement, the

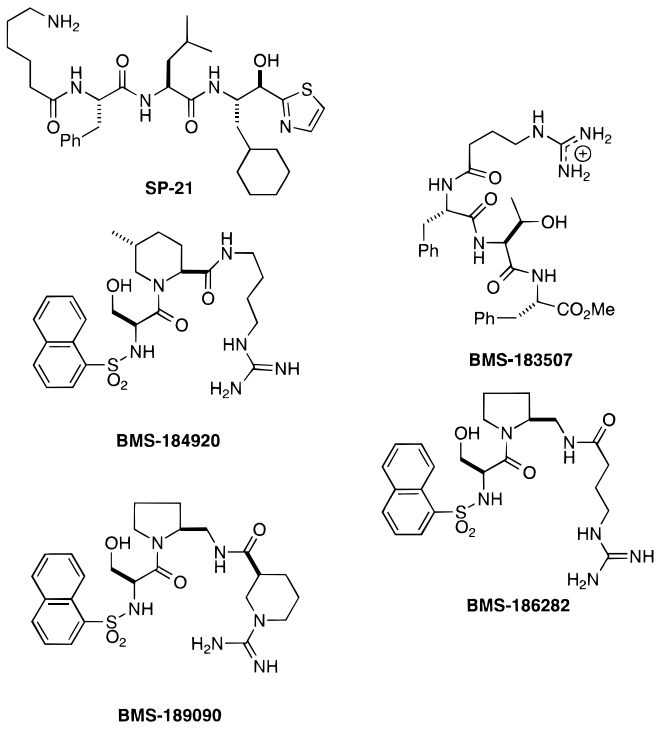


Figure SP*27. Structures of various inhibitors.

backbone of the tripeptide forms a parallel β-strand to the main chain of Gly-216. This is opposite to the orientation of substrate-like inhibitors such as **PPACK**. The guanidine group of the inhibitor makes electrostatic interaction in the S_1 specificity pocket similar, but not identical, to **PPACK**. The two phenyl groups make an edge to face interaction with each other. In addition, the hydroxyl group of the *allo*-threonine makes a hydrogen bond with the backbone NH of Gly-219.

Optimization of **BMS-183507**, based on the structure of its complex with thrombin and the structure of thrombin complexes with **PPACK**, **MD-805** and **NAPAP**, led to more potent and selective inhibitors. An early lead compound in this work was **BMS-184920**. A 7-fold increase in potency was obtained by replacement of the piperidine ring with a pyrrolidine ring and changing the direction of the amide bond linking the P_1 and P_2 groups; this resulted in **BMS-186282** (K_i = 78.9 nM). Optimization and conformational restriction of the guanidine side chain led to **BMS-189090** ($K_i = 3.64$ nM). This is one of the more potent and selective thrombin inhibitors reported. **BMS-189090** is a poor inhibitor of trypsin with a $K_i = 47$ μM. Structures of both **BMS-186282** and **BMS-18909** with thrombin have been reported (PDB entries: 1BMM, 1BMN)[210] (Figure SP*27).

Both **BMS-186282** and **BMS-18909** bind with the peptide bond facing the catalytic triad and the sulfonamide oxygens solvent exposed. (See Figure SP*28.) These inhibitors each form a pair of hydrogen bonds to Gly-216 in a manner typical of antiparallel β-strands and form a hydrogen bond to the carbonyl oxygen of Ser-214. The proline groups of the inhibitors occupy the S_2 subsite and the naphthalene groups occupy the P_4 subsite. The guanidine groups of both inhibitors make electrostatic interactions with the S_1 specificity pocket. While the bound conformations of the two inhibitors are quite similar

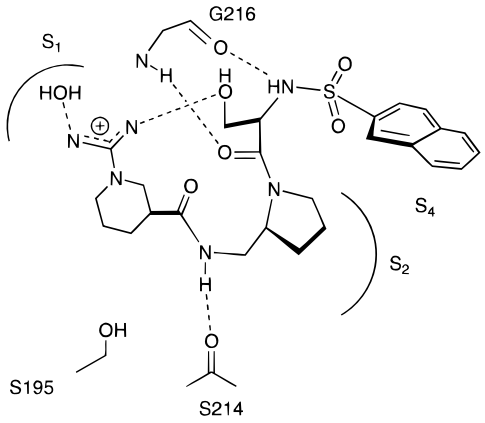


Figure SP*28. Observed binding mode of **BMS-18909**.

there are some significant differences in their conformations. The similarities include the antiparallel binding mode, the naphthalene group in the S_4 pocket, the proline groups in the S_2 pocket, and hydrogen binding with the backbone oxygen of Gly219.

The major differences in the inhibitor conformations include the hydrogen bonding of Glu192 to O17 of **BMS-186282**, in contrast to the **BMS-189090** structure where Glu-192 is turned away and hydrogen bonded to Thr-147. Due to the additional ring group in **BMS-189090**, the alkylguanidine moiety is shortened by 2 Å, relative to **BMS-186282**. Consequently, differences in their interaction in the S_1 specificity pocket are observed. Asp-189 binds directly to **BMS-186282**, but makes no contacts with **BMS-189090**. Interactions with the catalytic triad also differ, as well as those with solvent molecules. These two ternary complexes form a variety of close protein−ligand contacts. Experimental data indicate that **BMS-189090** is a 20-fold stronger inhibitor than **BMS-186282**. It is not entirely clear which variations are responsible for this difference in inhibition.

3. Hirulog Active-Site Hybrids

A number of thrombin inhibitors are modified proteins. One portion of the inhibitor binds to the hirudin "exosite", and another binds to the S_1 to S_4 substrate subsites. Many of these inhibitors thus interact with the catalytic triad as they bridge the S_1 and S_1' subsites. While a discussion of the protein−protein interactions is beyond the scope of this review, a short discussion of how these inhibitors interact with the catalytic triad is warranted.

The inhibitor **SP-22** (PDB entry: 1ABI; Figure SP*29)[211] has the Pro and D-Phe units present in **PPACK** which bind to the S_2 and S_4 subsites; however, it does so with a different geometry than does **PPACK**. The inhibitor has a β-amido arginine in the P_1 position. The presence of an additional methylene group alters the interaction of the arginine side chain with the S_1 specificity pocket and also the interactions of the D-Phe and Pro groups with the S_2 and S_4 subsites. The P_1 carbonyl is not in the oxyanion hole and the P_1' amide NH makes a hydrogen bond to His-57 which suggests that this group is neutral.

The inhibitor **SP-23** (PDB entry: 1DIT)[212] has a P_2-P_4 groups composed of L-Pro and L-Asp which is N acylated with (2-propyl)valeric acid. The *N*-acyl

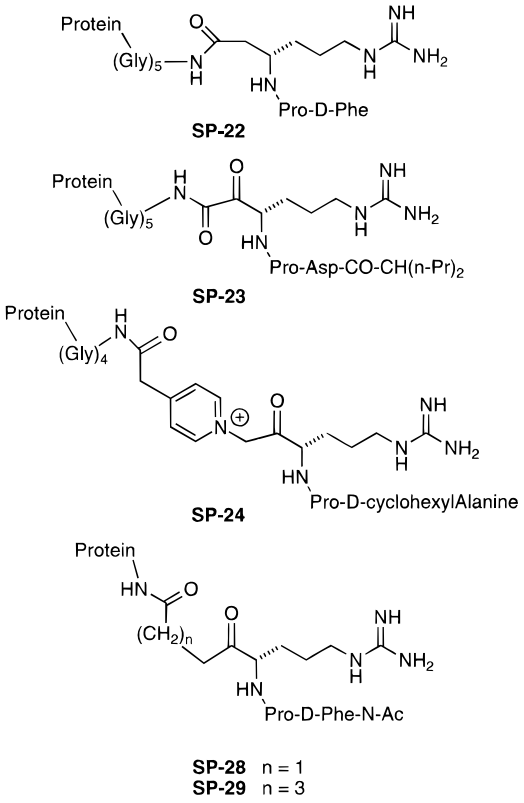


Figure SP*29. Partial structures of active-site hirulog hybrid inhibitors.

group makes VDW contacts in the S_4 subsite the proline makes VDW contacts with the S_2 subsite. The P_3 aspartate group is mostly solvent exposed and makes a hydrogen bond to the backbone NH of Gly-219. The arginine side chain interacts with the S_1 pocket in a fashion very similar to **PPACK**. The α-keto amide interacts with the catalytic triad in the same fashion as does cyclotheonamide A. The ketone carbonyl forms a covalent adduct with Ser-195 of the enzyme and the hemiketal oxygen occupies the oxyanion hole. The amide carbonyl forms a hydrogen bond to His-57 which is presumably protonated.

The inhibitor **SP-24** (PDB entry:1HBT)[213] has a D-cyclohexylalanine N-terminus which binds like the D-Phe of **PPACK**, the arginine and proline groups also bind like **PPACK**. The P_1' residue is a pyridinium group. The ketone carbonyl has undergone nucleophilic attack with the hemiketal oxygen in the oxyanion hole. The charged pyridinium group does not appear to make any intimate electrostatic interactions with the protein. It does make VDW contacts with the side chain of Trp-60, the hydrophobic part of the side chain of Glu-192 and the backbone of Gly-193.

An inhibitor related to **SP-24**, **SP-25**, has been reported to be a potent competitive inhibitor of thrombin with $K_i = 0.19$ nM[214] (Figure SP*30). To determine the influence of charge and size of the P_1' groups the related analogs **SP-26** ($K_i = 1.25$ nM) and **SP-27** ($K_i = 60.0$ nM) were prepared and evaluated. The presence of the positive charge contributes 1.1 kcal mol^{-1} to the stability of the complex and the VDW contacts from the aromatic ring contribute 2.2 kcal mol^{-1}. The positive charge may stabilize the complex by either stabilization of the oxyanion ad-

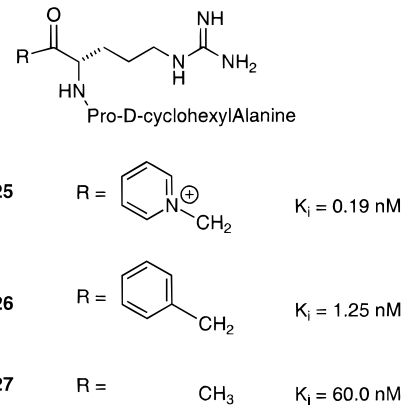


Figure SP*30. Structures of various inhibitors.

duct, or by an electron withdrawing effect on the carbonyl carbon or a combination of both.

One additional example involves unactivated ketones as linking groups, **SP-28** (PDB entry: 1IHS) and **SP-29** (PDB entry: 1IHT).[215] In these cases, the ketone does not form a covalent adduct with the enzyme. The ketone carbonyl oxygen is located in the oxyanion hole and the carbonyl carbon is in sub-VDW contact (2.8 Å) with the hydroxyl oxygen of Ser-195.

E. Elastase

1. Introduction

Human neutrophile elastase (HNE) (EC 3.4.21.37), human leukocyte elastase, is a serine protease of the chymotrypsin class. In general, elastases are one of the most destructive enzymes in the body having the ability to destroy connective tissue components. HNE is released from polymorphonuclear leukocytes in response to inflammatory stimuli and has been implicated in the development of various diseases such as emphysema, cystic fibrosis, and rheumatoid arthritis. As a result, studies directed toward small molecule inhibitors of HNE has been an active area of research.[216] While the structure of HNE has been determined by X-ray crystallography,[217,218] most structural studies on elastase inhibitors have been carried out with a related enzyme, porcine pancreatic elastase (PPE).[219]

Comparison of the structure of HNE (PDB entries: 1HNE, 1PPF, 1PPG) with thrombin reveals many differences. The most striking differences are in size, shape, and electronic nature of the S_1, S_2, and S_4 subsites. In contrast to thrombin, HNE has an open, solvent exposed S_2 subsite. The S_4 subsite is a shallow solvent exposed groove. The S_1 subsite is significantly different. Residues 190 and 216 are both valines and residue 192 is a phenylalanine. This makes the S_1 subsite a small well-defined hydrophobic pocket. The loop between Val-216 and Cys-220, which forms a disulfide link with Cys-191, has a different conformation than the corresponding loop in thrombin. As a result the substrate preferences for elastase are quite different. It has a preference for valine at P_1 and a preference for proline at P_2. The preference for proline at P_2 is presumably due to the bend it imparts along the peptide backbone.[220]

The structure of the complex between HNE and third domain of the turkey ovomucoid inhibitor

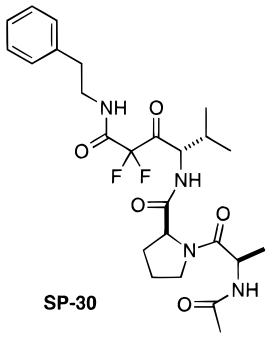


Figure SP*31. Structures of various inhibitors.

(TOMI), a natural protein-based inhibitor has been reported (PDB entry: 1PPF).[217] This structure is similar to the complex between chymotrypsin and eglin C since it resembles a Michaelis complex (Ser-195 OG−Leu I-18 C distance = 2.79 Å, C−O−N−CA improper torsion = 3.26°). This structure also shows the nature of the antiparallel hydrogen bonding between the inhibitor and Val-216 of HNE and the orientation of the inhibitor side chains from P5 to P_3'. This structure was used as starting point for HNE inhibitor design by the ZENECA group.[221]

2. Peptide-Based Inhibitors

The peptide based difluoro ketone inhibitor, **SP-30** (Figure SP*31), is a 100 nM competitive inhibitor of PPE at pH 7.8. The structure of its complex with PPE has been described (PDB entry: 4EST).[222] PPE shares about 40% homology with HNE. At the active site there are many similarities as well as significant differences. The S_1 subsite of PPE is larger than that of HNE, the substitution of Gly for Val at position 190 and Gln for Phe at position 192 are largely responsible for the increased size of the PPE S_1 subsite. In addition, the loop between Val-216 and Cys-220 have different conformation in these two enzymes. There are also small differences in the S_2-S_5 subsites of these two enzymes.

This complex shows features expected of peptidyl compound bound to a serine protease, specifically antiparallel hydrogen bonds between the backbone of the alanine residue of the inhibitor and the backbone of Val-216, and a long (3.42 Å) interaction between the NH of the valine-derived inhibitor residue and the backbone carbonyl of Ser-214. The difluoro ketone group has undergone nucleophilic attack by the alcohol oxygen of Ser-195 and the resulting hemiketal oxygen is in the oxyanion hole and makes close hydrogen bonds to the backbone NH's of Ser-195 and Gly-193 and the fluorine atoms close to His-57. The P_2' phenethyl side chain makes surface VDW contacts with the protein. The corresponding trifluoromethyl ketone analog of **SP-30** is a weaker inhibitor with K_i = 9.5 μM. Thus, the interactions of the prime side group with the enzyme contribute about 2.7 kcal mol^{-1} to the stability of the complex.

In pioneering work, the ICI (now ZENECA) group reported that a 2-benzoxazole group, α to a ketone, allows the carbonyl to react with the nucleophilic Ser-195 of elastase, resulting in potent enzyme inhibition. Inhibitor **SP-34** is a potent competitive inhibitor of HNE. In comparison, the methyl ketone **SP-32** is a

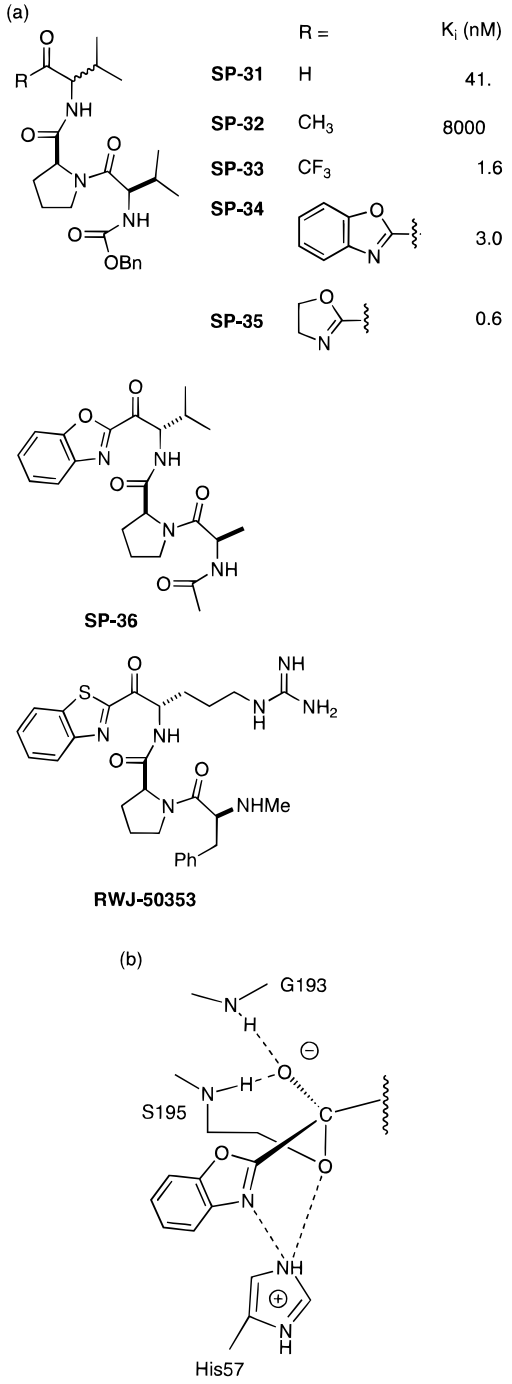


Figure SP*32. (a) Structures of various inhibitors. (b) Interaction of α-keto benzoxazole of **SP-36** with catalytic triad of PPE. The hemiketal oxygen (presumably the oxyanion) interacts with the oxyanion hole and the nitrogen of the benzoxazole hydrogen bonds with His-57 (presumably protonated).

much weaker inhibitor. Compound **SP-34** is about 2-fold less potent than the corresponding trifluoromethyl ketone (**SP-33**) and 14 times more potent than the corresponding aldehyde **SP-31**. Further SAR studies resulted in **SP-35** which is 2.7 times more potent than the trifluoromethyl ketone.[223]

The structure of the complex between **SP-36** and PPE has been reported.[223] (See Figure SP*32b.) The structure of the complex shows that the ketone carbonyl of **SP-36** has undergone nucleophilic attack by the alcohol oxygen of Ser-195 to form a hemiketal. The hemiketal oxygen is located in the oxyanion hole and 2.56 and 2.65 Å from the backbone NH's of Ser-195 and Gly-193, respectively. These close contacts suggest that the hemiketal exists as an oxyanion. The nitrogen atom of the benzoxazole is 2.81 Å from NE2 of His-57. This suggests that the histidine is doubly protonated. While the symmetric nature of the benzoxazole does not allow one to unambiguously establish if the nitrogen or oxygen is hydrogen bonding to the protonated imidazole, the assignment of the nitrogen in close contact with the imidazole was made on considerations of the relative basicity of the two heteroatoms. On the basis of this structure, it is quite likely that the ability of the heterocyclic ring to stabilize the positive charge on His-57 contributes significantly to the stability of the complex. The greater potency of **SP-35**, relative to **SP-34**, is consistent with this notion. Inhibitors **SP-34**, **SP-35**, and **SP-36**, and related compounds, were the first reported example of reversible, mechanism-derived inhibitors of serine proteases designed, and experimentally determined, to exploit the binding potential of the histidine, serine and oxyanion hole. Further SAR studies have determined that related heterocycles are also capable of being potent serine protease inhibitors.[224,225]

A related compound, **RWJ-50353**, has been reported to be a potent (K_i = 0.19 nM) thrombin inhibitor. The structure of its complex with thrombin has been reported.[226] The different size of nitrogen and sulfur allows for the determination that the nitrogen atom of the heterocycle forms the hydrogen bond to His-57.

Trifluoroacetyl peptides provide another class of potent, reversible inhibitors of elastase with K_i values between 10 nM to 1 μM.[227] The crystal structures of a number of these inhibitors show that they bind as "retro-peptides", and that they have three different types of binding orientations.[228]

The related peptides **SP-37** (PDB entry: 1ELB), **SP-38** (PDB entry: 1ELA), and **SP-39** (PDB entry: 1ELC) are inhibitors of PPE with K_i's that only differ by a factor of about 6 (Figure SP*33, Table SP*1). The structure of the complex between each of these inhibitors with PPE show that they each bind differently to the enzyme. The most potent inhibitor, **SP-37**, forms only one hydrogen bond to the protein between the NH_2 of Gln-200 and the carbonyl of the leucine of the inhibitor. The leucine side chain makes tight VDW contacts in the S_1 subsite, the aromatic ring of the ISO (*p*-isopropylaniline) group makes VDW contacts with the hydrophobic atoms of the side chain of Arg-226 and the lysine side chain of the inhibitor. The trifluoromethyl group makes contacts with both Ser-195 and His-57. Inhibitor **SP-38** binds much differently, the trifluoromethyl group now makes VDW contacts in the S_1 subsite and the ISO group now makes VDW contacts with the floor of the S_4 subsite. There are three protein−ligand hydrogen bonds, between the NH_2 of Gln-200 and the trifluoroacetyl carbonyl, between the proline carbonyl oxygen and the backbone NH of Val-224, and between the ISO NH and the backbone carbonyl of Val-224. The position occupied by the trifluoromethyl group of **SP-37** is now occupied by an acetate molecule. Compound **SP-39** binds in yet a different manner,

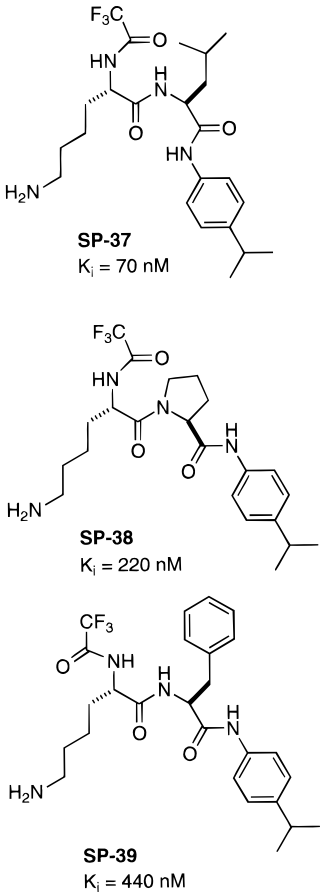


Figure SP*33. Structures of various inhibitors.

Table SP*1. Functional Groups of TFA-Dipeptide-Anilide Inhibitors That Are Found in Each of the PPE Subsites

inhibitor	OAH	S1	S2	S3	S4	SA	SB
SP-37	TFA	Leu	Lys			ISO	
SP-38		TFA	Lys	Pro	ISO		
SP-39	TFA	Leu	Lys				ISO

[a] Legend: OAH = "oxyanion hole"; SA and SB = different binding sites for ISO group.

which is somewhat similar to that of **SP-37**. There is one hydrogen bond between the protein and the inhibitor, between the side-chain carbonyl of Gln-200 and the ISO amide NH. The phenylalanine side chain occupies the S_1 subsite, and the CF_3 group interacts with Ser-195 and His-57 as did **SP-37**. The hydrophobic portion of the ISO group makes VDW surface contact with the loop between Ser-225 and Gly-228; this contact buries a carbonyl oxygen of the protein. For each of these three inhibitors, the lysine side chain makes weak VDW contacts in the S_2 subsite.

3. Nonpeptidic Reversible Inhibitors

While peptidic compounds can be very potent elastase inhibitors, they typically have properties which make them unsuitable as bioavailable drug candidates. The Zeneca group has reported a series of papers which deal with this problem.[220,221,229,230]

Modeling studies, based on the structures of available complexes of PPE and HNE with peptide-based inhibitors, were a starting point for this work. The

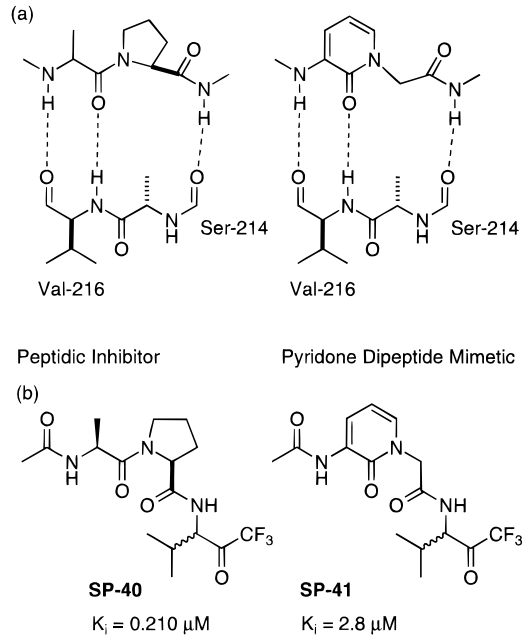


Figure SP*34. (a) Protein−ligand hydrogen-bond interactions of a peptidic inhibitor which are mimicked by pyridone dipeptide mimetic. (b) Structures of various inhibitors.

antiparallel hydrogen-bonding interactions between the enzyme and inhibitors characteristic of serine proteases was targeted for replacement with a nonpeptidic framework. A pyridone dipeptide mimetic was designed as first generation compounds. (See Figure SP*34.) Initial results showed a 13-fold loss in potency.[221,229]

Further design considerations suggested functionalized analogs of **SP-41** which would make better VDW contacts with the S_2 and S_4 HNE subsites. This resulted in a number of potent HNE inhibitors, including **SP-42** for which a crystal structure of its complex with PPE has been reported (PDB entry: 1EAS; Figure SP*35).[231,232] The structure of the complex of **SP-42** with PPE shows that the inhibitor binds as modeled and that the pyridone unit acts as a dipeptide mimetic. The phenyl group makes additional VDW contacts with the enzyme in the S_2 subsite. In addition, many of these compounds are potent HNE inhibitors and are orally active.

Structures of the related inhibitors **SP-45** (PDB entry: 1EAT) and **SP-46** (PDB entry: 1EAU) have also been reported and bind to PPE similar to **SP-42**. A pyridopyrimidine analog **SP-47** has been reported to be a subnanomolar inhibitor of HNE. A comparison of analogs **SP-48** and **SP-49** shows that boronate **SP-49** is about 8-fold more potent than TFK **SP-48**.[233]

These and related studies[234,235] by the Zeneca group provide a nice example of using information about protein structure to design novel, nonpeptidic inhibitors of an important enzyme target.

F. Summary: Reversible Serine Protease Inhibitors

From structural studies on a number of potent inhibitors of serine proteases which form a covalent bond to the active site serine, a significant trend is observed. Inhibitors containing a neutral group **G**

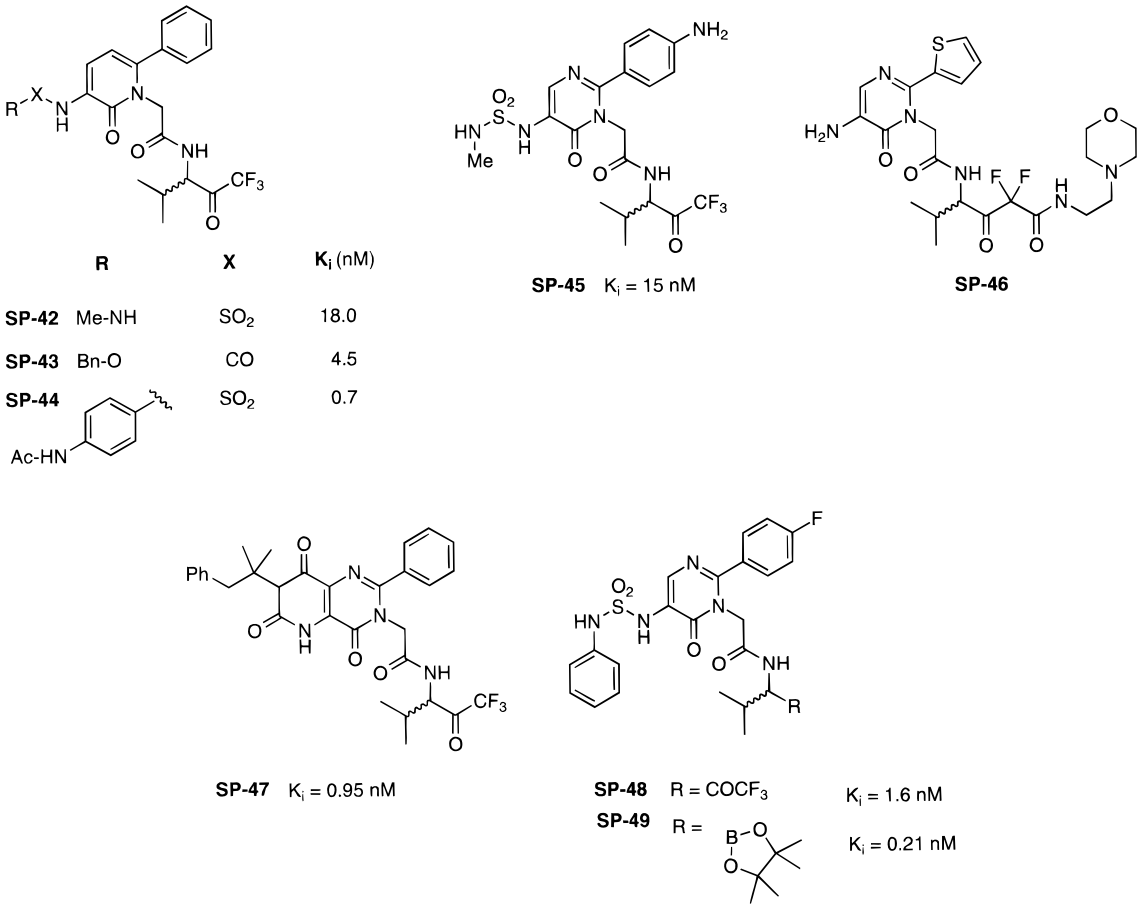


Figure SP*35. Structures of various inhibitors.

that form a covalent bond to the active-site serine result in the conversion of an enzyme resting state where only the aspartate is charged to a negatively charged adduct (**G**$^-$) and a positively charged histidine. On the basis of the high potency of these types of inhibitors this is a highly favorable process. It is likely that the serine proteases have evolved to stabilize this sort of electronic arrangement. The protonation states of the inhibited complexes have been determined by a number of experimental methods. They include neutron diffraction for phosphates, NMR studies for trifluoromethyl ketone and boronic acids, and examination of close heteroatom−heteroatom contacts using X-ray diffraction.

Some of these classes of compounds have been shown to be transition-state analogs on the basis of kinetic studies. Kinetic isotope studies have also shown that C−N bond cleavage is not very significant in the rate-determining step for amide hydrolysis.[134] The available evidence suggests that serine proteases achieve catalysis by stabilizing the formation of a charged tetrahedral intermediate. It is likely that serine proteases have evolved to stabilize the conversion of a neutral ligand (substrate), nucleophile (Ser-195), and base (His-57) into an electronic environment where the base becomes protonated and a negative charge resides in the oxyanion hole. The potent inhibitor classes in Figure SP*36 are able to mimic this situation and take advantage of the evolved ability of theses enzymes to stabilize a transition state.

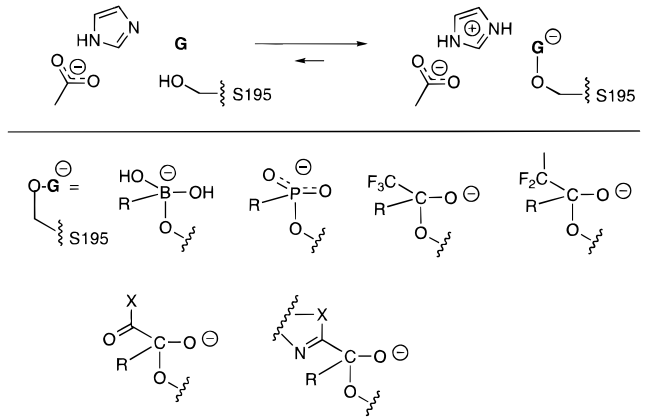


Figure SP*36. Groups **G** which form a covalent bond to Ser-195 to generate an anion and result in a positively charged His-57. Ligands which contain these groups are typically potent serine protease inhibitors.

G. Irreversible Inhibitors

There are a number of irreversible inhibitors of serine protease and an overview of these inhibitors has been reviewed.[220] While an extensive discussion of these classes of inhibitors is outside of the scope of this review, a brief discussion of some of these compounds is appropriate.

The β-lactam **SP-50** is an irreversible inhibitor of elastase. The structure of its inhibited complex with PPE has been reported.[236] The structure of the inhibited complex shows that it has formed covalent bonds with both Ser-195 and His-57. (See Figure SP*37.) The alkylation of His-57 is responsible for

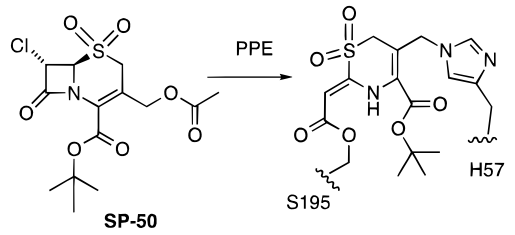


Figure SP*37. β-Lactam **SP-50** is an irreversible inhibitor of PPE which alkylates NE2 of His-57. Alkylation of His-57 is not essential for inhibition of serine proteases by β-lactams, acylation of Ser-195 and slow hydrolysis of the acyl enzyme adduct is sufficient for effective inhibition.

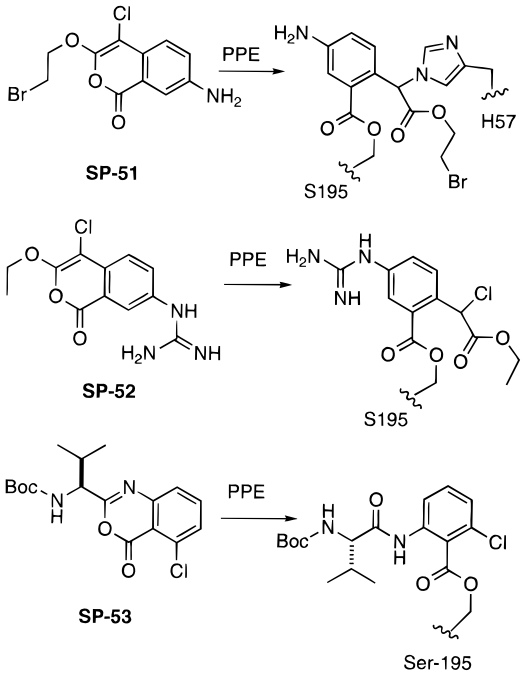


Figure SP*38. (a) Isocoumarin **SP-51** alkylates His-57 of PPE. (b) Isocoumarin **SP-52** acylates PPE but does not alkylate His-57. Slow hydrolysis of the acyl enzyme adduct is responsible for PPE inhibition. (c) Benzoxinone **SP-53** inhibits PPE, it acylates Ser-195 and slow hydrolysis of the acyl enzyme adduct is responsible for responsible for inhibition.

the irreversible nature of the complex. Other β-lactam inhibitors of elastase do not alkylate His-57 and are not irreversible. However, many of these β-lactams are effective inhibitors because the acyl enzyme complex is very slowly hydrolyzed to regenerate active enzyme.[237]

Another class of irreversible inhibitors for PPE for which the structure of the inhibited complex has been solved are isocoumarins and benzoxinones. (See Figure SP*38.) Isocoumarin **SP-51** forms covalent bonds to both Ser-195 and His-57. The related inhibitor **SP-52**, forms only an ester bond with Ser-195 and does not alkylate His-57. The inhibitory potency of **SP-52** is likely due to very slow hydrolysis of the acyl ester enzyme complex. The structure of PPE complex inhibited by the benzoxinone inhibitor **SP-53** has been reported.[238] This compound only acylates Ser-195, again a very slow off rate for hydrolysis is responsible for inhibition. Compounds **SP-52** and **SP-53** are neither reversible or irreversible inhibitors, but rather are very enzyme substrates

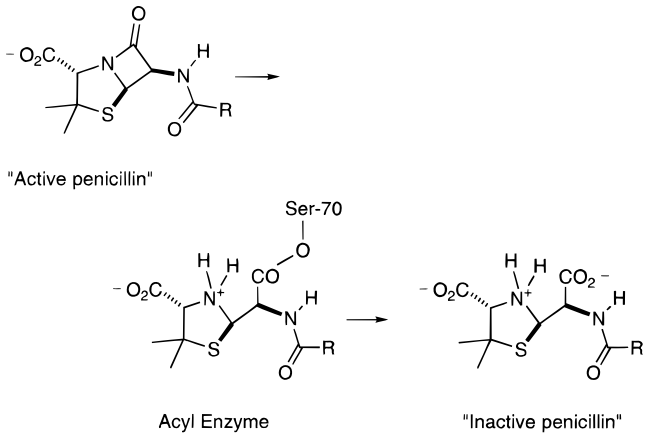


Figure SP*39. β-Lactamases hydrolyze and inactivate penicillins via an intermediate acyl enzyme complex.

which act as inhibitors due to their slow rate of turnover.

The structure of these complexes provide structural evidence for ultimate product of the inhibited complexes. It also allows one to propose a mechanism for inhibition but, in many cases, provides relatively little insight into the molecular recognition of the ligand prior to irreversible chemistry. The structure of the complex between chymotrypsin and the peptide-based irreversible inhibitor, (2*S*)-*N*-acetyl-L-alanyl-L-phenylalanyl-α-chloroethane, showed that a substantial conformational change in the complex occurred after the irreversible step.[182]

H. Enzyme Related to Serine Proteases

1. β-Lactamases

β-Lactam antibiotics are inhibitors of a class of transpeptidase enzymes that are responsible for bacterial cell wall formation.[239] The β-lactamases, bacterial enzymes that hydrolyze and destroy β-lactam antibiotics, are mechanistically related to serine proteases. (See Figure SP*39.) These plasmid encoded enzymes result in bacteria which are resistant to many β-lactam antibiotics. The structure of TEM-1, a representative member of the class A β-lactamases, has been described (PDB entry: 1BTL).[240]

Many residues in the class A β-lactamase active site are conserved in both Gram-positive and Gram-negative species of bacteria. These residues include Ser-70, Lys-73, Ser-130, Asn-132, Glu-66, Asn-170, Lys-234, Ser-235, and Arg-244. The hydroxyl group of Ser-70 is the nucleophilic group which attacks the β-lactam carbonyl to form an acyl enzyme intermediate, the acyl enzyme is then rapidly hydrolyzed by an enzyme-activated water molecule (see Figure SP*40).[241−244] Thus, the class A β-lactamase enzymes are mechanistically related to the serine proteases.

While these enzymes use a serine hydroxyl group to attack an amide (lactam) carbonyl, the structure of the β-lactamase active site is unrelated to any known serine protease. The hydroxyl oxygen of Ser-70 is 2.9 Å from NZ of Lys-73. Lys-73, whose protonation state is unknown, is not solvent accessible. It appears likely that NZ of Lys-73 acts as a general base and general acid in the catalytic mechanism, analogous to His-57 in the chymotrypsin class

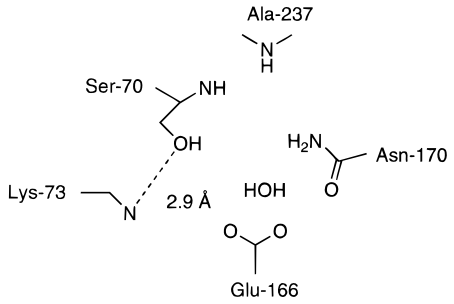


Figure SP*40. Active-site residues of TEM-1 β-lactamase. Ser-70 is the active-site nucleophile and Lys-73 presumably acts as a general base and general acid, the oxyanion hole is formed by the backbone NH's of Ser-70 and Ala-237. A structural water molecule, bound between Asn-170 and Glu-166 hydrolyses the acyl enzyme adduct.

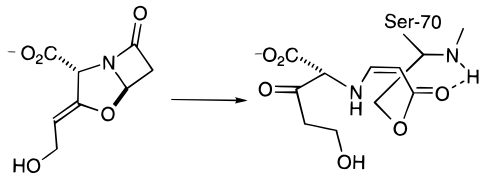


Figure SP*41. Clavulanic acid (**SP-54**) inactivates β-lactamases by being converted into the vinylogous urethane acyl enzyme **SP-55**.

of proteases. The enzyme's oxyanion hole is formed from the backbone NH's of Ser-70 and Ala-237. A structural water molecule is observed between the polar side chains of Glu-166 and Asn-170. This water molecule, activated by the general base Glu-166, is likely to be the nucleophile which hydrolyzes the acyl enzyme intermediate.

Clavulanic acid (**SP-54**), is a natural product which is a mechanism-based inhibitor of class A β-lactamases.[245] The structure of the inhibited complex has been determined (PDB entry: 1BLC).[246] This structure shows that **SP-54** forms an acyl enzyme complex, with Ser-70, which has subsequently been converted into a vinylogous urethane **SP-55**. (See Figure SP*41.) The inhibition of the enzyme by **SP-54** is due to the inability of the enzyme to efficiently hydrolyze the degradation product vinylogous urethane **SP-55**.

Boronic acid **SP-56** is a potent (K_i = 110 nM) reversible inhibitor of TEM-1 β-lactamase.[247] The structure of its complex shows a tetrahedral boronate species.[248] (See Figure SP*42.) One of the oxygens on boron interacts with the oxyanion hole and the other oxygen interacts with the polar side chains of Glu-166 and Asn-170. The latter results in displacing the structural water molecule which acts as the nucleophile for hydrolysis of the acyl enzyme complex.

On the basis of the structure and mechanism of the TEM-1 β-lactamases, **SP-57** was designed as an inhibitor. It was proposed that **SP-57** would acylate the enzyme, but it would resist deacylation by virtue of the hydroxymethyl group. This group was designed to displace the structural water responsible for the deacylation step. In fact, **SP-57** inactivated the TEM-1 β-lactamase rapidly in a process which was saturable. It acylates the enzyme, but resists deacylation.[249] In contrast, penicillinic acid **SP-58** is a substrate. (See Figure SP*43.)

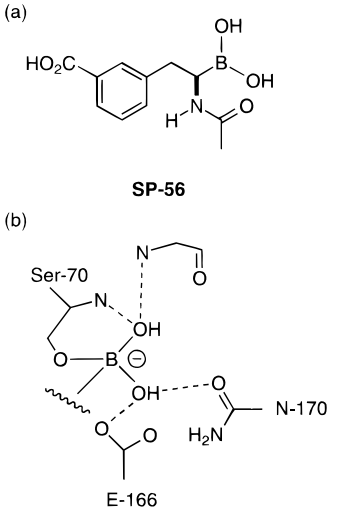


Figure SP*42. (a) Boronic acid **SP-56** is a reversible inhibitor of TEM-1 β-lactamase, and (b) it forms a stable type I adduct with the enzyme.

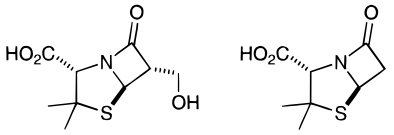


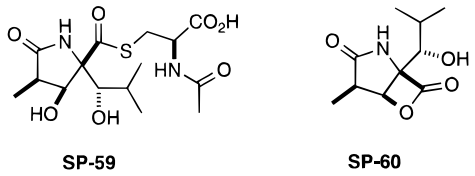
Figure SP*43. Compound **SP-57** is a β-lactamase inhibitor while **SP-58** is a substrate. The hydroxyl group of **SP-57** interacts with the structural water molecule which acts as the nucleophile for acyl enzyme hydrolysis.

The crystal structure of the inhibited complex has been reported (PDB entry: 1TEM).[250] This inhibited complex is, as expected, the acyl enzyme complex. In this structure, the structural water is not displaced by the hydroxyl group of the inhibitor, rather the hydroxyl group of the inhibitor forms a hydrogen bond to this water molecule. The presence of this hydroxymethyl group may sterically hinder the approach of the structural water molecule to the ester carbonyl of the acyl enzyme.

2. Proteosomes: Threonine Proteases

Most proteins in eukaryotic cells are degraded by a soluble, ATP-dependent pathway present in both the nucleus and cytosol. The key proteolytic component in this pathway is the 20S proteosome. This pathway is known to degrade many critical regulatory proteins that must be rapidly destroyed for normal growth and metabolism.[251] In addition to its "housekeeping function" in cell protein degradation, the proteosome has been implicated in a variety of disease states ranging from immune diseases to cancer.[252−254] The natural product lactacystin (**SP-59**), and its β-lactone analog (**SP-60**), are irreversible inhibitors of the 20S proteosome (Figure SP*44). Lactacystin inhibits cell cycle progression and induces neurite outgrowth in a murine neuroblastoma cell line.[255]

The crystal structure of the 20S proteosome from the archaebacterium *Thermoplasma acidophilium* has been determined at 3.4 Å resolution (PDB entry: 1PMA).[256] This large (673 kD) protease complex consists of 14 copies of two different sub-



SP-59 SP-60

Figure SP*44. Structures of proteosome inhibitors.

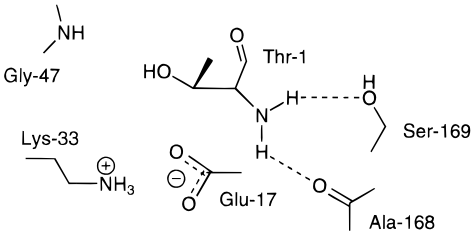


Figure SP*45. Active-site residues of the *T. acidophilium* proteosome, a threonine protease.

units, α and β, and has 14 active sites. The active site is in the β subunit. (The eukaryotic proteosome is more complex with multiple α and β type subunits.) What is unique about this structure is that it is a threonine protease, and the active-site threonine is the N-terminal residue (i.e., Thr-β1). (See Figure SP*45.) Lactacystin (**SP-59**) and **SP-60** have both been found to covalently modify this conserved N-terminal threonine in the mammalian proteosome.[255] A 4.0 Å structure of the *T. acidophilium* proteosome in complex with a tripeptide aldehyde shows the C-terminal of inhibitor in close contact with Thr-β1. The active site suggests a catalytic tetrad consisting of Thr-1 OG, Thr-1 N, Lys-33 NZ (presumed protonated), and Glu-17. On the basis of the low resolution of the peptide aldehyde complex it was proposed that the backbone NH of Gly-47 may be the oxyanion hole. Further structural and mechanistic studies will likely be needed to further clarify the protease mechanism. The 20S proteosome represents a unique protease enzyme which could potentially be a future drug-design target.[257]

V. Inhibitors of Cysteine Proteases

A. Introduction

Cysteine protease are a distinct mechanistic class of protease enzymes which take advantage of the nucleophilcity of a thiolate to achieve amide hydrolysis. To date, there are three structural families of these enzymes which have been characterized: the papain class, the ICE class and the picornovirus 3C-proteases class.

Review articles concerning cysteine proteases have appeared.[258−261]

B. Papain Class

The papain class of cysteine proteases is the most well-characterized member of these proteases. The active site of papain, like chymotrypsin, contains a catalytic triad (Cys-25, His-159, and Asn-175) and an oxyanion hole (Cys-25 and Gln-19). In papain, an asparagine residue is the "third member" of the triad. (See Figure CP*1.) Like subtilisin, the active site nucleophilic residue lies at the N-terminus of an

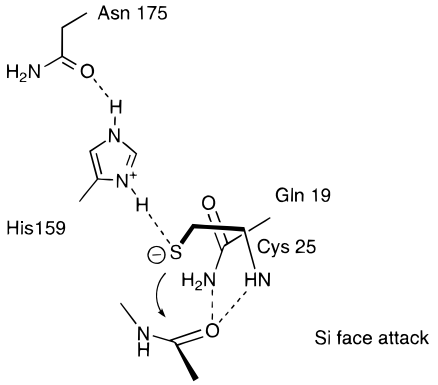


Figure CP*1. Catalytic triad of papain. The topology of the active site results in Si face addition to peptide substrates.

α-helix, and the side chain of a glutamine form part of the oxyanion hole. Examination of the structure of papain, and substrate-like inhibitors, shows that the three-dimensional orientations of the catalytic triad and oxyanion hole require that nucleophilic attack on the carbonyl takes place from the Si face. This is in contrast to all other well-characterized classes of serine and cysteine proteases. The resting state of papain appears to involve a negatively charged cysteine and a positively charged histidine.[262,263]

The protease mechanism of papain is related to the protease mechanism of chymotrypsin. The nucleophile (thiolate anion) attacks the carbonyl of the substrate to give a tetrahedral intermediate. The tetrahedral intermediate breaks down with loss of the amine product to give an acyl enzyme intermediate. The acyl enzyme intermediate then reacts with water to give the product acid and free enzyme. For papain-catalyzed hydrolysis of benzoyl-L-arginine amide the $^{14}N/^{15}N$ kinetic isotope observed was considered as evidence that there is significant C−N bond cleavage occurring in the transition state. This suggests that breakdown of the tetrahedral intermediate is rate limiting.[264] In contrast, the kinetic isotope effect for chymotrypsin-catalyzed hydrolysis is much smaller and has been interpreted as evidence for a tetrahedral intermediate in which both formation and breakdown contribute to the rate-limiting step.[134] Thus, kinetic evidence supports a "later" transition state for hydrolysis by papain which is closer to the acyl intermediate.

Cathepsin B is a lysosomal enzyme which catalyzes normal protein degradation but has also been associated with several pathophysiological conditions such as inflammation, tumor metastasis, bone resorption, muscular dystrophy, and myocardial infarction. Therefore, there is a need for selective inhibitors of cathepsin B which are effective in vivo.[265] The plant enzyme papain serves as an excellent model for cathepsin B; the amino acid sequences of these enzymes show a very high degree of homology in the active-site regions, and it has been concluded that the overall protein folding patterns and catalytic mechanisms are the same.[266,267]

The structure of the natural product leupeptin (**SP-2**) in its complex with papain has been reported (PDB entry: 1POP; see Figure CP*2).[268] The sulfur

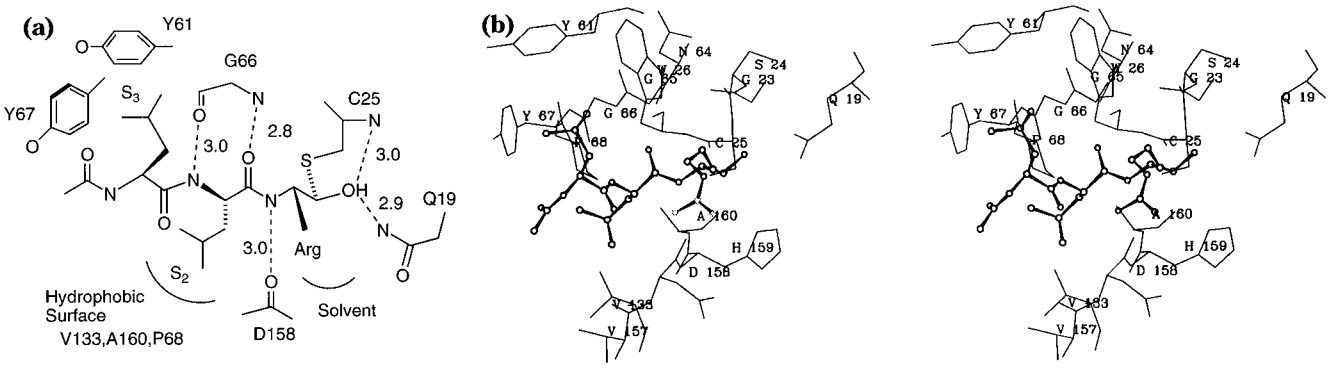
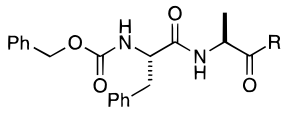


Figure CP*2. (a) Interactions of leupeptin with papain. The S_1 subsite is solvent exposed and the S_2 and S_3 binding sites are hydrophobic. The P_2 backbone of the inhibitor forms hydrogen bonds with the backbone of Gly-66. (b) Stereoview of leupeptin−papain complex.



		K_i (μM)
CP-1	R = H	0.021
CP-2	R = CH_3	31.
CP-3	R = CF_3	
	Isomer A	300.
	Isomer B	470.

Figure CP*3. Structures and affinities of various ligands.

of Cys-25 has attacked the aldehyde to form a hemithioacetal. The oxygen atom forms hydrogen bonds in the oxyanion hole with the backbone NH of Cys-25 and the side chain of Gln-19. Only one orientation for this oxygen was observed. Papain does not have a well-defined S_1 subsite and the amidine group of the arginine side chain is solvent exposed. There are three backbone to backbone hydrogen bonds between the inhibitor and the enzyme (two to Gly-166 and one to Asp-158). The S_2 and S_3 subsites of papain are both hydrophobic and the leucine side chains of the inhibitor make tight VDW contacts with both subsites. The S_2 subsite is a hydrophobic surface defined by Val-133, Ala-160, and Pro-68. The S_3 subsite is formed by two aromatic groups, Tyr-61 and Tyr-67.

A ^{13}C NMR experiment has been reported for the complex between papain and (*N*-acetylphenylalanyl)-glycinal. In this experiment two diastereomeric tetrahedral hemithioacetal adducts were observed.[269] This situation is similar to the crystallographic observation of two diastereomeric hemiacetal adducts formed between peptidyl aldehydes and serine proteases in the chymotrypsin class.[143,144]

Leupeptin is also a potent competitive inhibitor of cathepsin B (K_i = 4.1 nM). The shorter peptide aldehyde **CP-1** (R = H) is also a cathepsin B inhibitor (K_i = 21 nM). (See Figure CP*3.) The analogous ketone (**CP-2**) is a much weaker inhibitor (K_i = 31 μM). Surprisingly, the analogous trifluoromethyl ketones, for example **CP-3**, are very weak inhibitors of cathepsin B (K_i = 300 μM); they are more than 10^5-fold less potent than the analogous aldehyde.[270] This is in sharp contrast to their respective abilities to inhibit members of the chymotrypsin class of serine proteases.

For trifluoromethyl ketones in aqueous solution the equilibrium strongly favors the hydrate. An evaluation of competitive hydration and hemithioketal formation in a model system (aldehydes, methyl ketones, and trifluoromethyl ketones) led to a correlation of cathepsin B inhibition constants (K_i) and model system equilibria between ketone (aldehyde), hydrate, and thioketal (thioacetal).[271] The poor inhibition of cysteine proteases by trifluoromethyl ketones reflects the chemically unfavorable equilibrium between the hydrate and the hemithioketal. The potent inhibition of serine proteases by trifluoromethyl ketones is the anomaly. In spite of an unfavorable chemical equilibrium, extremely favorable protein ligand (oxyanion) interactions make the equilibrium between the hydrate and the serine protease adduct extremely favorable.

One speculation (see Figure CP*4) is that for the serine proteases, going from the enzyme resting state, where only the aspartate is charged, to the TFK adduct, in which in addition to the aspartate the histidine is charged and the hemiketal is an oxyanion, is a highly favorable process. It is likely that these enzymes have evolved to stabilize this sort of electronic arrangement. For cathepsin B, the resting state involves a charged histidine and charged cysteine. If addition to the TFK carbonyl by cysteine proteases behaved analogously, one would now have a charged histidine and a charged oxyanion. In this case, the distance between the opposing charges will increase as the reaction proceeds. It appears that this equilibrium is not quite as favorable as in the serine protease case.

Other studies have reported that peptide nitrile **CP-4** is a reversible, competitive inhibitor of papain (K_i = 0.73 μM).[272] This compound is not hydrolyzed by papain. In contrast, nitriles have not been found to be good inhibitors of serine proteases.[273,274] Nitrile **CP-5** inhibits papain (K_i = 380 μM), but is not a substrate. The analogous aldehyde **CP-6** is a more potent inhibitor of papain (K_i = 25 μM). ^{13}C NMR experiments have reported that the inhibition of papain by nitrile **CP-5** results in the enzyme bound thioimidate ester adduct **CP-7**[158,275] (Figure CP*5). The protonation state of His-159 was not determined. However, since thioimidate formation requires protonation, it is likely that His-159 is neutral in the complex. Thus, it appears that these adducts are structural analogs of the acyl enzyme intermediate

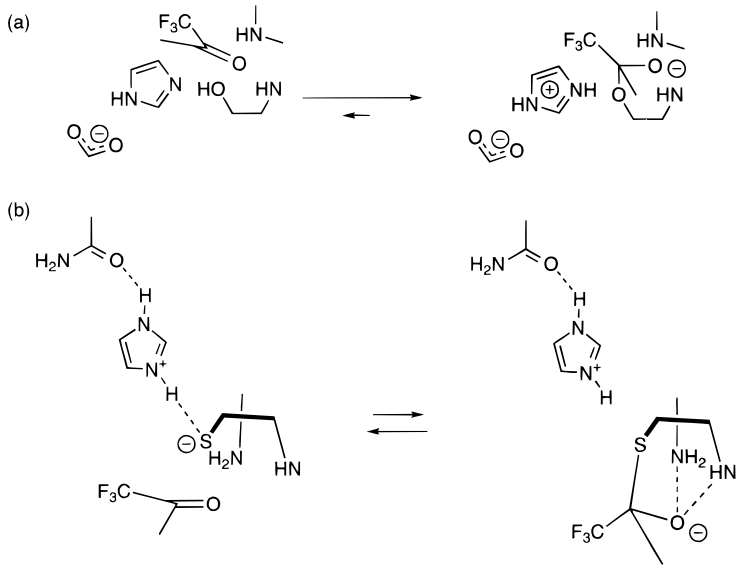


Figure CP*4. (a) The covalent interaction between a TMK inhibitor and a serine protease produces a highly favorable electrostatic interaction. (b) The covalent interaction between a TMK inhibitor and a cysteine protease is not favorable.

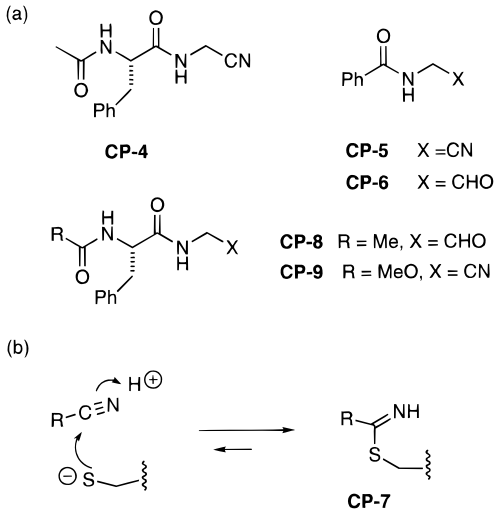


Figure CP*5. (a) Structures of various inhibitors and (b) cysteine proteases reacting with peptidyl nitriles to form an enzyme bound thioimidate ester adduct **CP-7**.

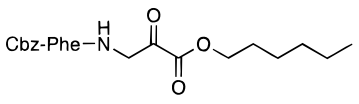
formed during peptide hydrolysis. For several *p*-nitroanilide substrates and their analogous inhibitors, a correlation between K_i and k_{cat}/K_m was observed. That finding, together with the fact that the pH dependence on K_i also parallels that of k_{cat}/K_m, suggests that peptide nitrile inhibitors have considerable transition-state character.[158]

An in-depth study of papain inhibition by aldehyde **CP-8** (K_i = 39.8 nM) and nitrile **CP-9** (K_i = 840 nM) has been described.[276] Papain oxyanion hole mutants Q19A and Q19E were found to be less potently inhibited by **CP-9**, while much smaller effects were observed with aldehyde **CP-8**. The influence of these mutations on pH dependency of inhibition by **CP-9** suggests that inhibition of papain by peptide nitriles is a process closer to that of substrate hydrolysis than is the inhibition by the corresponding peptide aldehydes. Since the papain nitrile adducts are structural analogs of the acyl intermediate, and not a tetrahedral adduct, this suggests that peptide hydrolysis catalyzed by papain has a "late" transition state which resembles the acyl adduct. This observation is consistent with the kinetic evidence that the transition state for hydrolysis by papain is "late" and closer to the acyl intermediate.

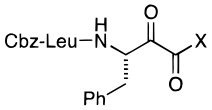
What has emerged from these studies is that the transition state that papain has evolved to stabilize is different than the transition state that chymotrypsin has evolved to stabilize. Papain stabilizes a later, more neutral transition state, while chymotrypsin stabilizes an earlier, more charged transition state. Therefore, inhibitors which attempt to mimic transition states along the hydrolysis pathway will have different effects on these very different enzymes. In fact, this appears to be the case. Trifluoromethyl ketones which mimic a charged transition state are potent inhibitors of chymotrypsin but are poor inhibitors of papain. Peptide nitriles, which mimic a neutral acyl intermediate, are potent inhibitors of papain but poor inhibitors of chymotrypsin.

The inhibition of cathepsin B and papain by a number of peptidyl α-keto esters, α-keto amides, α-diketones, and α-keto acids has been reported.[277] These compounds are slow binding inhibitors. The proposed mechanism of inhibition involves formation of a hemiketal by active site thiol addition to the ketone carbonyl group of the inhibitor that is closer to the N-terminus. The electron-withdrawing effect of the second carbonyl likely stabilizes the hemithioketal. The most potent of these inhibitors was **CP-10** (K_i (papain) = 800 nM; K_i (cathepsin B) = 150 nM). (See Figure CP*6.)

Calpains are calcium-dependent cysteine proteases which are widely distributed in mammalian cells and are involved in a number of diverse biological functions.[278] There are two distinct classes of calpains: calpain I requires micromolar concentrations of calcium for activation while calpain II requires millimolar concentrations of calcium. No structures of any member of the calpain family has been reported. The calpains appear to be related to the papain superfamily; however, plant papain and the human calpain families have since undergone significant evolutionary divergence.[279]

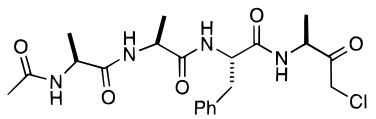


CP-10



		K_i (μM)		
		Cal I	Cal II	Papain
CP-11	X = CO_2Et	1.8	0.40	75.
CP-12	X = CONHEt	0.20	0.039	45.
CP-13	X = CO_2H	0.0085	0.0057	7.

Figure CP*6. Structures and affinities of various inhibitors.



CP-14

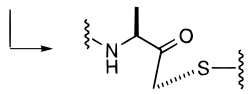


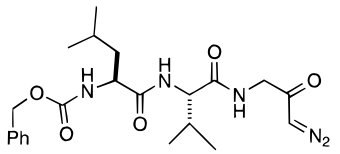
Figure CP*7. α-Chloro ketones such as **CP-14** irreversibly alkylate the active-site cysteine.

Peptide α-keto esters, α-keto amides, and α-keto acids have been described as inhibitors of calpains and papain. While these compounds can be potent, low nanomolar inhibitors of calpains they are much poorer inhibitors of papain.[280]

Peptidic α-chloro ketones are irreversible inhibitors of cysteine proteases. These inhibitors form a carbon−sulfur bond between the enzyme and the methylene group containing the chlorine. The structures of a number of covalent complexes between members of the papain class and peptidic α-chloro ketones have been reported.[281] For **CP-14** (PDB entry: 1PAD; Figure CP*7) the peptide portion of the inhibitor binds to papain in a fashion similar to leupeptin. The backbone of the phenylalanine side chain of **CP-14** makes antiparallel hydrogen bonds to the backbone of Gly-66. The hydrophobic phenylalanine side chain makes extensive VDW contacts in the hydrophobic S_2 subsite. The remainder of the inhibitor does not make extensive contacts with the protein. The ketone group is in the vicinity of the oxyanion hole but does not make any hydrogen bonds to the protein.

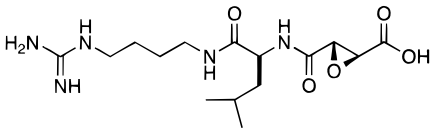
The chemical behavior of α-chloro ketone inhibitors with papain is quite different than their behavior with serine proteases. For serine proteases, it is the active-site histidine that is alkylated, not the active-site serine. In the papain complexes the active site histidine does not interact with the inhibitor.

Glycyl endopeptidase, a plant enzyme, is a member of the papain family which has specificity for cleavage C-terminal to glycine residues. In papain and cathepsin B, the S_1 subsite is mostly solvent exposed and does not have an influence on substrate specificity. The covalent complex between glycyl endopeptidase and the irreversible inhibitor **CP-15** has been reported (PDB entry: 1GEC; Figure CP*8).[282] Inhibitor **CP-15**, an α-diazo ketone, alkylates the active-site cysteine to give an α-thio ketone product. The



CP-15

Figure CP*8. α-Diazo ketones such as **CP-15** irreversibly alkylate the active-site cysteine.



E-64

Figure CP*9. The natural product **E-64** is an irreversible inhibitor of cysteine proteases in the papain class.

structure of its complex provides insight into the enzyme's selectivity for substrates with a glycine in the P_1 position. Substitution of the glycines at residues 23 and 65 in papain, with glutamic acid and arginine, respectively, in glycyl endopeptidase introduce two groups into solvent. These two groups now form an extremely small S_1 subsite which can only accommodate a glycine residue at P_1. Thus, the glycine selectivity results from the exclusion of substrates with larger P_1 residues.

This product between the α-diazo ketone **CP-15** and the enzyme gave the same S-alkylated product which would have been formed from an α-chloro ketone inhibitor. Other irreversible inhibitors of cysteine proteases which are proposed to have similar chemical behavior are α-fluoro ketones,[283] acyloxymethyl ketones,[271] and ketomethylsulfonium salts.[284]

The natural product **E-64** is a potent and irreversible inhibitor of a number of cysteine proteases, including actinidin, papain, ficin, and cathepsins B and L.R.[259] The structure of the complexes between **E-64** and both papain[285] and actinidin (PDB entry: 1AEC; Figure CP*9)[286] have been reported. The actinidin complex has been solved at 1.86 Å resolution and shows that the epoxide group has alkylated the active site cysteine (Cys-25). This alkylation occurs with inversion of configuration; the resulting hydroxyl group is solvent exposed and does not interact with the protein. This bond formation is responsible for the irreversible nature of the inhibition. In the inhibited complex, a number of favorable interactions between the inhibitor and the protein are noted. The carboxylate makes several electrostatic interactions with the protein. One carboxylate oxygen makes hydrogen-bonding interactions to the oxyanion hole, with the backbone NH of Cys-25 (2.89 Å) and the side chain NH_2 of Gln-19 (2.86 Å). The other carboxylate oxygen makes a hydrogen bond to ND1 of the active-site His-162 (2.75 Å), which is presumably protonated, and another to a water molecule. The peptide backbone of **E-64** binds to the protein in a reverse direction than does a peptide aldehyde, that is, the amino to carboxy directions are opposite. However, the peptide backbone of the leucine residue of **E-64** still makes hydrogen bonds to the protein that are similar to the hydrogen bonds between leupeptin and papain. The leucine side chain makes VDW contacts with the protein in the

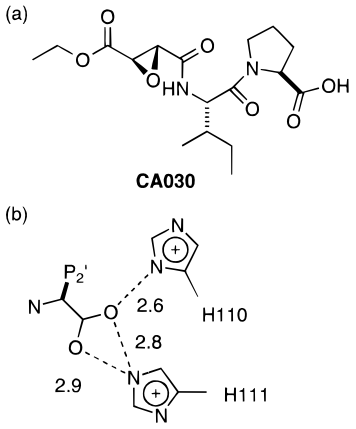


Figure CP*10. (a) Structure of cathepsin B inhibitor **CA030**. (b) Cathepsin B has a preference to cleave the last two residues from a peptide chain. The structure of the complex between CA030 and cathepsin B shows interactions between the carboxylate of the inhibitor and His-110 and His-111 of the protein, this type of interaction is likely responsible for the sequence selectivity of cathepsin B.

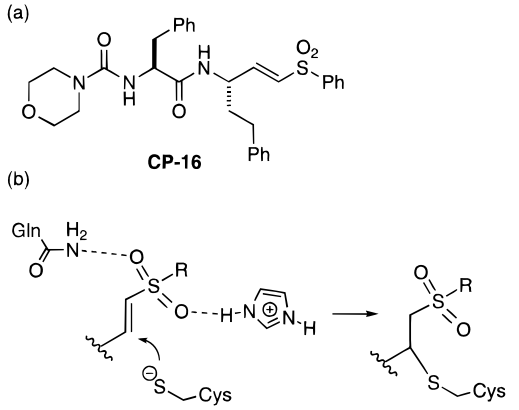


Figure CP*11. (a) Vinyl sulfones are selective, irreversible inhibitors of cysteine proteases in the papain class. (b) Proposed interactions between a vinyl sulfone and active-site histidine and glutamine.

hydrophobic S_2 subsite, the guanidine side chain is solvent exposed. The mechanism of inactivation by **E-64** likely involves noncovalent interactions of the inhibitor with the protein in a manner similar to that seen in the complex, followed by nucleophilic attack by the thiolate on the epoxide to irreversibly form a carbon sulfur bond. The **E-64** papain complex is very similar. Structures of other inhibitors related to **E-64** irreversibly bound to papain have been reported (PDB entries: 1PE6[287] and 1PPP[288]). These inhibitors bind to papain in the same fashion that **E-64** does.

The structure of the complex between a related inhibitor, **CA030**, and a related protein, cathepsin B, shows a significantly different structure (see Figure CP*10).[289] One difference between cathepsin B and other members of the papain family is its preference to cleave the last two residues from a peptide chain, that is, it has carboxydipeptidyl activity. The complex between **CA030** and cathepsin B (PDB entry: 1CSB) shows that alkylation of the thiolate by the epoxide has occurred. However, it appears that the opposite stereochemical face of the epoxide has been attacked (an unambiguous stereochemical assignment of the epoxide adduct was not achieved). This results in the peptide chain in this complex running in the same direction that would be expected for a substrate. In this complex the peptide residues of **CA030** act as P_1' and P_2' groups and thus identify the S_1' and S_2' subsites of the enzyme. In cathepsin B a large insertion loop (residues 106−124) blocks the primed side of the active site cleft. The residues His-110 and His-111 interact with the C-terminal carboxyl group of the inhibitor. The syn-face of the carboxyl group interacts with NE2 of His-111 and the anti-lone pair of one of the carboxyl groups interacts with ND1 of His-110. An interaction of this type is likely responsible for the substrate selectivity of cathepsin B. The isoleucine side chain makes VDW contacts with the protein in the vicinity of Val-176 and defines the S_1' subsite. In this complex a hydrogen bond between the succinyl amide carbonyl and the amide NH of the side chain of Gln-23, which forms part of the oxyanion hole, is observed. The proline ring and the ethyl group of the ester make VDW contacts with the protein.

By superimposing the structures of the complexes of papain with peptide aldehydes (PDB entry: 1POP) and covalently bound chloromethyl ketones (PDB entry: 1PAD) on the structure of cathepsin bound to **CA030**, models for substrate binding to members of the papain class of enzymes can be constructed. These models illustrate substrate binding from P_3 to P_2'.[289]

Mechanistic studies on the inactivation of papain by epoxysuccinyl inhibitors have been reported.[290] These studies indicate that a carbonyl group α to the epoxide which is attacked by the thiolate is necessary for good activity. These studies also indicate that protonation of the epoxide during the inactivation process likely occurs from solvent and not from His-159. This study also suggests that inhibitor potency depends on the affinity of the epoxide for the enzyme (K_d) prior to irreversible chemistry.

Vinyl sulfones were designed as mechanism-based cysteine protease inhibitors. (See Figure CP*11.) The design process was based on enzyme mechanism and active site geometry. It was proposed that vinyl sulfones would hydrogen bond to the side chain NH_2 of the oxyanion hole glutamine and the protonated active-site histidine. This recognition process would serve to orient the vinyl sulfone in the active site and make it vulnerable toward nucleophilic attack by the active-site thiolate. Vinyl sulfones are not very reactive toward nucleophiles, and it was proposed that enzyme activation would be needed for reaction to occur.[291]

A number of vinyl sulfones related to **CP-16** were prepared and evaluated as inhibitors of various cysteine proteases. Compound **CP-16** was shown to irreversibly inhibit cathepsins B, L, S, and O2 and cruzain. The compound also showed different potencies toward these different enzymes. These compounds did not react with dithiothreitol nor did an analog of **CP-16**, designed to be complementary toward HNE, inhibit the serine protease HNE. It was thus proposed that peptide-based vinyl sulfones are effectively inert in the absence of the papain superfamily's catalytic machinery. To date, no struc-

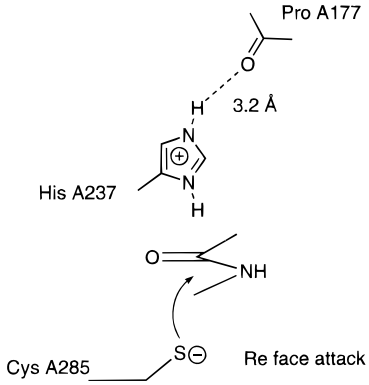


Figure CP*12. The active-site residues of ICE show that this is a structurally distinct class of cysteine protease enzymes.

ture of a vinyl sulfone bound to a cysteine protease has been reported.

C. ICE Class

Interleukin-1β-converting enzyme (ICE) is a cysteine protease responsible for the production of interleukin-1β (IL-1β) in monocytes.[292] This enzyme thus has a role in the pathogenesis of inflammatory diseases such as rheumatoid arthritis. ICE is homologous to CED-3, the gene product required for apoptosis in *Caenorhabditis elegans*. At least one related human enzyme, CPP32, appears to play a role in mammalian apoptosis.[292] The structure of ICE (PDB entry: 1ICE) has been determined by two different groups[293,294] and is a new class of cysteine protease enzymes. The function of ICE is to cleave a precursor of IL-1β into the biologically active form of IL-1β, it cleaves at an Asp-Ala site. ICE shows selectivity for substrates with an aspartate in the P_1 position.

The structure of ICE shows that it is unrelated to the papain family. ICE is a heterodimer, the active site contains residues from both monomers. The structure of ICE has been solved in its complex with the tetrapeptide aldehyde inhibitor **CP-17** ($K_i = 0.76$ nM).[294] The active site has a catalytic dyad, consisting of Cys-A285 and His-A237. In the complex, NE2 of His-A237 forms a long hydrogen bond to the backbone carbonyl of Pro-A177. In the complex, the distance between ND1 of His-A237 and the nucleophilic sulfur of Cys-A285 is 5.3 Å. The hemithioacetal carbon resulting from thiolate addition to the aldehyde is located between the Cys-A285 and His-A237. This is a unique structural feature not seen in any other serine or cysteine protease whose structure has been determined. The hemithioacetal oxygen makes a hydrogen bond (2.65 Å) to ND1 of His-A237. The oxyanion hole, which appears to consist of the backbone NH's of Cys-A285 and Gly-A238 is unoccupied. On the basis of the structure of this adduct and the location of the oxyanion hole, it would appear that the thiolate nucleophile of ICE would attack normal substrates from the Re stereochemical face of the scissile amide carbonyl. (See Figure CP*12.)

The structure of the ICE complex with **CP-17** also provides some insight into the preference of ICE for substrates with an aspartate P_1 group. The S_1 subsite of ICE contains two arginine side chains. The carboxylate P_1 group of **CP-17** makes several strong hydrogen bonds with the S_1 groups (see Figure CP*13). Each syn lone pair of the carboxylate makes a hydrogen bond to an NH of the guanidine side chain of Arg-A179 (2.7 and 2.9 Å), in addition one carboxylate oxygen interacts with the side chain of Arg-B341 and the other oxygen interacts with the amide NH of the side chain of Gln-A283. Thus, the S_1 subsite of ICE provides strong electrostatic complementarity to the side chain of an aspartate residue. The backbone NH of the P_1 residue makes a hydrogen bond (3.0 Å) to the backbone carbonyl oxygen of Ser-B341.

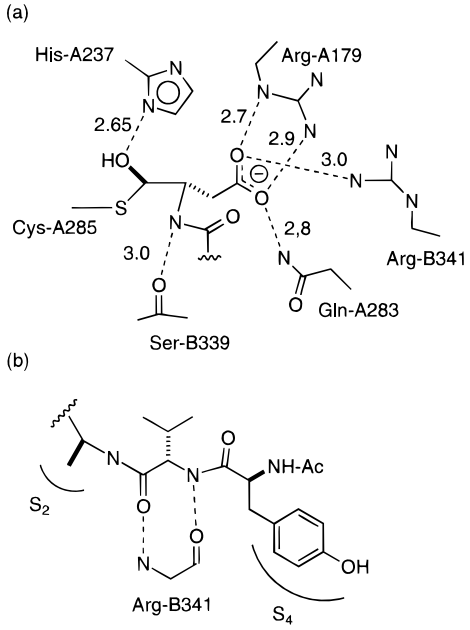


Figure CP*13. (a) Interactions of P_1 side chain of **CP-17** with S_1 specificity pocket of ICE. (b) Interactions of P_2-P_4 groups of **CP-17** with ICE.

The remainder of the inhibitor makes less strong interactions with the protein. The P_2 alanine side chain sits in a small hydrophobic S_2 pocket formed by the side chains of Val-B338 and Trp-B340. The backbone of the P_3 valine residue makes antiparallel hydrogen bonds to the backbone of Arg-B341, the P_3 side chain is mostly solvent exposed only making very weak VDW contacts with the protein. The phenol P_4 side chain sits in the deep S_4 cavity of ICE and makes extensive VDW contacts. In addition, modeling studies have suggested that a substrate would also make extensive contacts of the "prime side" (P_1' to P_4').[294]

Some SAR studies on peptidic inhibitors of ICE have been reported. Reversible inhibitors of ICE include peptide aldehydes,[295–298] nitriles,[299] and ketones.[300–303] The higher potency of ketone **CP-20** compared to **CP-19**[301] is consistent with the hydrophobic $(CH_2)_5Ph$ group of **CP-20** making hydrophobic contacts with the large primed side of ICE. As was the case for the papain class of proteases, a difluoro ketone **CP-21** was found to be a poor inhibitor of ICE ($K_i = 6\ \mu M$).[302] (See Figure CP*14.)

The search for potent, selective, nonpeptidic inhibitors of ICE is likely to be an important area of future

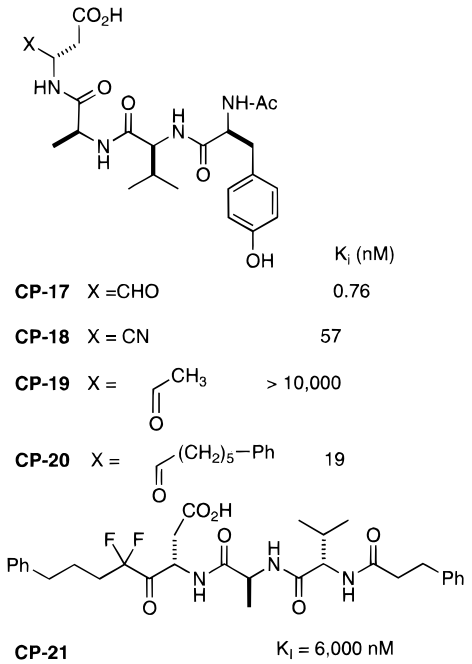


Figure CP*14. Structure and affinities for ICE of various ligands.

research. The structure of ICE should play an important role in the discovery of these compounds.

D. Picornovirus 3C-Proteases

The picornoviruses are an RNA family of viruses that include hepatitis A virus, poliovirus, and the rhinovirus responsible for the common cold. The maturation of these viruses require cleavage of polyproteins; these cleavages are performed by the virally encoded 3C-protease. These 3C-proteases are cysteine proteases.[304,305] The active-site cysteine is present in a Gly-Xxx-Cys-Gly-Gly motif that is related to the conserved Gly-Xxx-Ser-Gly-Gly motif found in the chymotrypsin family of serine proteases. This led to two independent proposals that the 3C-proteases could be homologously aligned with the chymotrypsin like proteases.[306,307] In these proposals the active site Ser-195 of chymotrypsin aligned with the conserved cysteines of the viral proteases, His-57 of chymotrypsin aligned with a conserved viral histidine. These homology studies also predicted that these viral proteases have conserved not only active-site geometries but also the typical double β-barrel protein fold. The solved structures of rhinovirus-14 3C-protease (RVP)[308] and hepatitis A virus 3C-protease (HAVP)[309] proved these predictions to be essentially correct. However, detailed comparisons of RVP and chymotrypsin reveal subtle differences. These differences demonstrate how minor changes in one part of a protein need to be compensated for in other parts of the protein in order to achieve catalysis (i.e., efficient molecular recognition of a transition state).

An experiment has been described in which the active-site serine-195 of rat anionic trypsin was replaced with a cysteine by site-directed mutagenesis. This mutation results in a single atom substitution (S for O) in a large protein. The k_{cat} of trypsin S195C was reduced by a factor of 6.5×10^5 relative to that

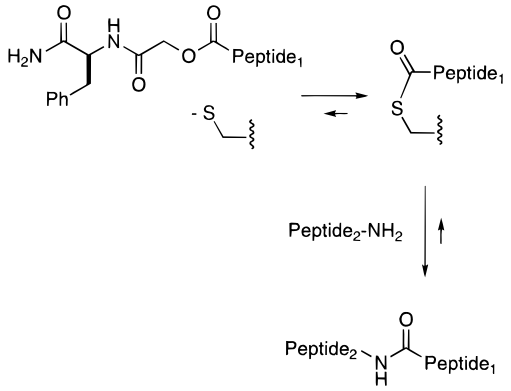


Figure CP*15. Subtiligase, a double mutant (S221C and P225A) of the serine protease subtilisin is an efficient protein ligase. Subtiligase reacts with appropriate ester substrates to give an acyl enzyme (thioester) intermediate, this enzyme bound intermediate reacts with peptides with free N-termini to form larger peptides. Thus, this mutant enzyme catalyzes the reverse reaction that the native protease does.

of trypsin while the K_m values were essentially unaffected.[310] This result suggests that the mutant protein, while still capable of binding its substrates, is deficient in its ability to tightly bind the transition state for peptide cleavage. The crystal structure of this mutant enzyme has been reported.[311] Structure analysis revealed only subtle changes in the conformation of the mutant protein. The sulfur atom is larger than the oxygen it replaced but does not appear to be sterically hindered. Due to a bond rotation the sulfur atom appears to be able to form a hydrogen bond to the backbone NH of Gly-193 which is part of the oxyanion hole. This occludes the oxyanion hole and may effect the ability of the mutant to achieve efficient catalysis. In addition, longer bonds to sulfur, relative to oxygen, and different bond angles probably do not allow efficient stabilization of the transition state for peptide hydrolysis. This result shows how minor changes can have a dramatic effect on the binding of a protein to its ligand, in this case a transition state.

For the subtilisin class of serine proteases single atom substitution of an oxygen, with a sulfur, in the active-site residue Ser-221 results in a new protein which has an altered function. With appropriate substrates this protein will now catalyze peptide bond formation.[312] Likewise, selenosubtilisin also acts as an acyl transferase (PDB entry: 1SEL).[313,314] The double subtilisin mutant (S221C and P225A) is a highly efficient peptide ligase called subtiligase.[315] Again, this is an example where very small changes in a protein have a dramatic influence on its function.

Subtiligase is capable of catalyzing the conversion of peptide esters into an acyl enzyme intermediate. (See Figure CP*15.) This acyl enzyme intermediate is slowly hydrolyzed by water. Subtiligase is, however, efficient at catalyzing the conversion of this acyl enzyme intermediate and other peptides with free amino N-termini into larger peptides and is a very poor catalyst for peptide hydrolysis. This designed, mutant enzyme catalyzes the reverse reaction of its natural counterpart. This data implies that for a cysteine protease to be an efficient protease it must be capable of converting a peptide substrate to an

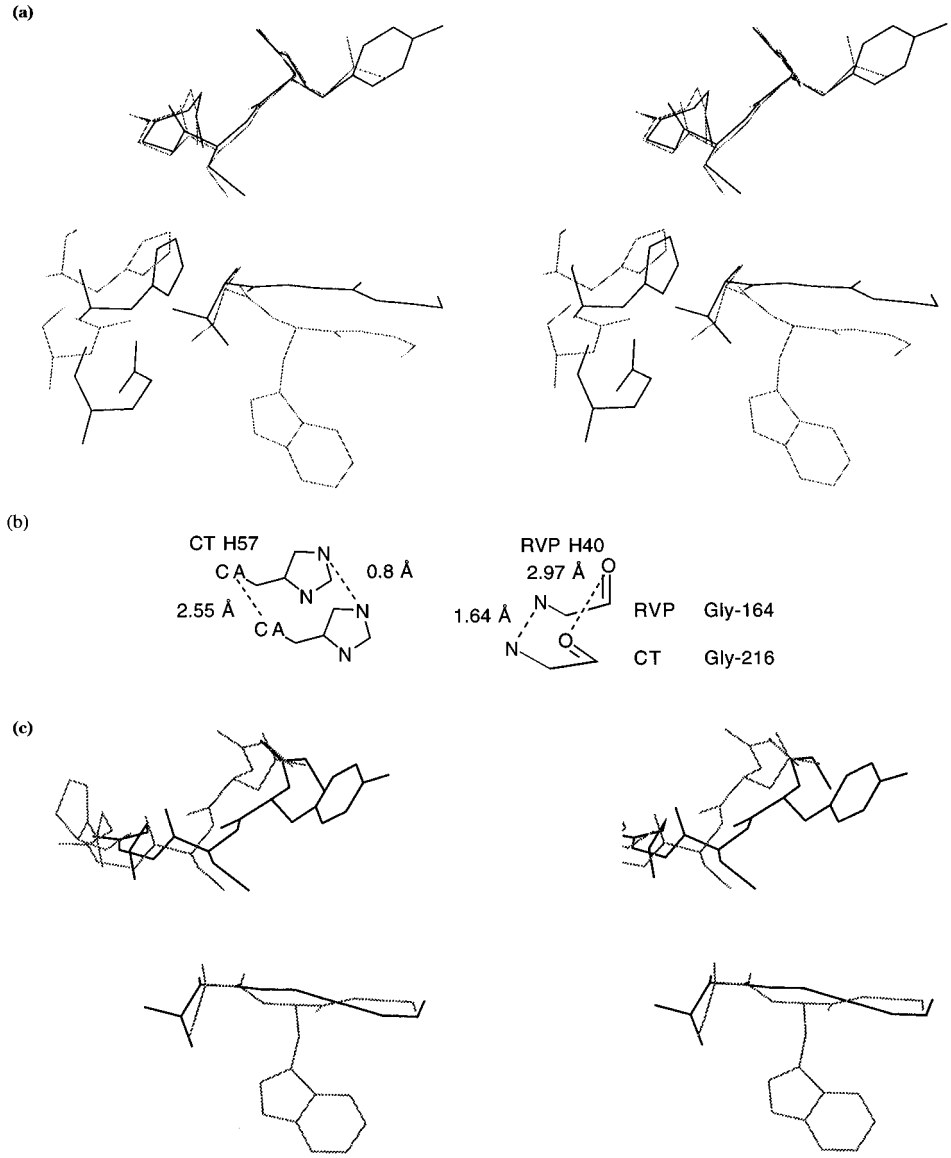


Figure CP*16. Picornoviral 3C-proteases have a cysteine nucleophile in a serine protease fold. (a) Alignment of RVP (bold) with chymotrypsin (gray) by superimposing the backbone atoms of the oxyanion hole of each protein (RVP residues 145−149 and CT residues 193−197). (b) In this alignment key catalytic atoms align, but other atoms do not. The P_2-P_4 substrate binding sites do not superimpose well with this alignment. (c) Alignment of RVP (bold) with chymotrypsin (gray) by superimposing the backbone atoms of the P_2-P_4 substrate binding sites of each protein (RVP residues 162−164 and CT residues 214−216). This alignment clearly shows the altered location of the oxyanion hole of each enzyme relative to the substrate binding site.

enzyme-bound acyl intermediate and hydrolyzing the acyl intermediate. The structure of subtiligase has been reported.[316]

With these mutational studies as a background it becomes quite interesting to observed how nature has evolved to perform the conversion of serine to a cysteine, in the context of a protein related to chymotrypsin, and create a catalytically useful enzyme. The structures of RVP and HAVP provide insight into this issue.

Superimposing the backbone atoms of residues 193 to 197 (active-site serine = 195) of chymotrypsin (CT) on the backbone atoms of residues 145 to 149 (active-site cysteine = 147) of RVP (see Figure CP*16) provides a nearly perfect overlap of each protein's oxyanion hole. With this alignment, the nucleophilic oxygen of chymotrypsin and sulfur of RVP nearly overlap (0.8 Å apart). However, as one examines other important regions of the proteins there are significant differences in the location of key groups relative to the active site nucleophile. For the active-site histidines (CT residue 57; RVP residue 40) the NE2 nitrogens are nearly coincident being only 0.8 Å apart; however, the CA atoms of these residues are now 2.55 Å apart. For RVP, the third member of the catalytic triad is a glutamic acid (E71), not an aspartate as in CT (D102). The relative position of these two groups are quite different in these two proteins, the CG and CD atoms of these two groups are 2.99 Å apart when the oxyanion hole are aligned. In RVP, another significant difference is that the carboxylate uses its anti-lone pairs to interact with its histidine partner, while in CT the syn lone pairs are used. For HAVP, it is likely that a tyrosine (Tyr-143) may be the third member of the triad. Also, the relative positions of the key substrate binding resi-

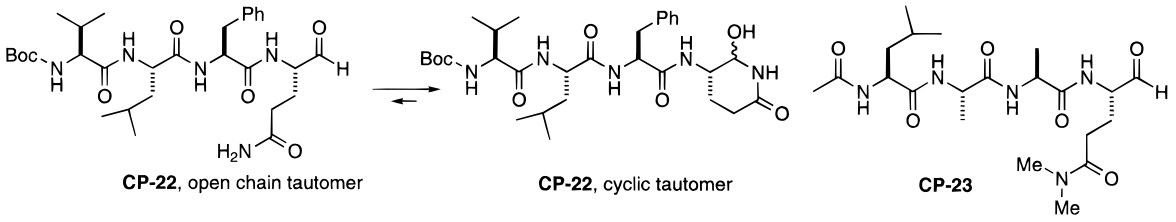


Figure CP*17. Structures of peptide aldehyde inhibitors, note that **CP-22** exists mostly as the cyclic tautomer.

dues Trp-215 and Gly-216 of CT and the homologous Gly-163 and Gly-164 of RVP are positioned differently relative to the nucleophilic atom. The RVP substrate binding site is actually closer to the RVP oxyanion hole (CA C147−CA G164 distance = 9.66 Å) than the CT substrate binding site is to the CT oxyanion hole (CA S195−CA G216 distance = 9.80 Å). Close examination of RVP and CT demonstrates how a nucleophilic cysteine can be accommodated into the active site of a chymotrypsin like protease. To accommodate a larger sulfur atom, longer sulfur carbon bonds and different bond angles around sulfur, and presumably a different rate-determining step, small subtle changes in the different regions of the active site are needed to achieve catalysis.

An alternative alignment, superimposing substrate binding domains (RVP backbone atoms 145−149 superimposed on CT backbone atoms 193−197) shows an altered location of the oxyanion hole in each of these protein. Since a difficult step for a cysteine protease to catalyze is the conversion of a substrate to an acyl enzyme intermediate, it is possible that the altered position of the oxyanion hole in RVP may assist in catalyzing this difficult conversion.

A site-directed mutagenesis studies in which the active-site cysteine of poliovirus 3C-protease was replaced with a serine has been reported. This study has shown that the C147S mutant is capable of processing precursor polyproteins; however, it is much less efficient than the wild-type enzyme. The C147A mutant was not capable of performing this processing.[317] No quantitation of the level of enzyme activity in the C147S mutant was reported.

An important lesson to be learned from the structure of RVP is that small changes in one part of a protein can be compensated for by other small changes in parts of the protein up to 10 to 15 Å away. Since catalysis involves effective molecular recognition of a transition state, a corollary to the RVP story would be that molecular recognition of ligands by proteins could also be dramatically effected by equally subtle features.

Since the structures of RVP and HAVP were reported, the following reports of inhibitors have appeared in the literature. Glutamine-derived inhibitors, such as **CP-22** (IC_{50} = 0.6 μM; Figure CP*17) are inhibitors of RVP.[318] The NMR spectrum of **CP-22** in CD_3OD lacked an aldehyde peak, and it was suggested that it exists mainly in the cyclic tautomer form. It is likely that observed IC_{50} reflects this unfavorable equilibrium in solution and that the open-chain tautomer is actually a more potent inhibitor than the measured IC_{50} value. In addition **CP-22** shows weak activity against rhinovirus-14 in cell culture. To date, no structures of RVP in complex with a peptidic ligand have been reported. However,

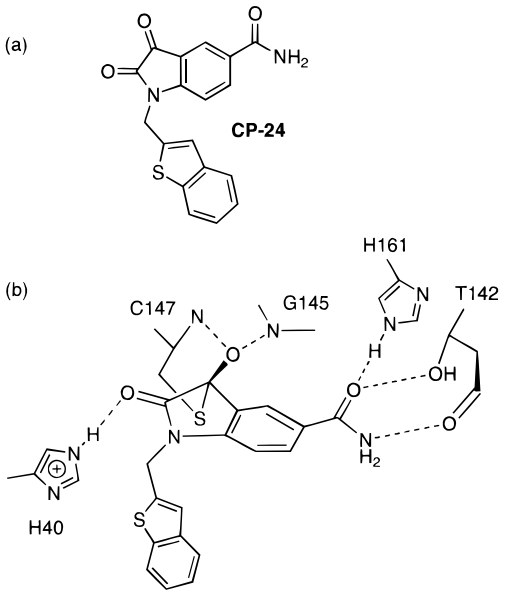


Figure CP*18. (a) Structure of the potent, reversible RVP inhibitor **CP-24** and (b) key interactions of **CP-24** with active site of RVP.

the Lilly group has mentioned that NMR studies are being used to study ligands related to **CP-22**.[319] Peptide aldehyde inhibitors of HAVP have also been reported recently.[320] Aldehyde **CP-23** was found to be a reversible, slow binding inhibitor of HAVP with K_i = 42 nM. This inhibitor showed 60-fold less activity against RVP. NMR studies showed that **CP-23** formed a covalent hemithioacetal adduct with HAVP. These NMR studies show only a single cross peak in the ^{13}C HMQC spectrum. This suggests that only a single diastereomeric adduct is formed.

In a nice example of structure-based inhibitor design, the Agouron group reported that isatin **CP-24** was a potent competitive inhibitor of RVP (K_i = 3.0 nM).[321] The structure of the complex between **CP-24** and rhinovirus-14 3C-protease (RVP-2) has been reported. (See Figure CP*18.) The reactive ketone carbonyl of the isatin has undergone nucleophilic attack by the thiolate of Cys-147, to give a hemithioketal. The hemithioketal oxygen, whose protonation state is unknown, forms hydrogen bonds to the backbone NH's of Cys-147 and Gly-145 which constitute the oxyanion hole. The lactam carbonyl of the isatin ring forms a hydrogen bond to NE2 of His-40, which is presumed protonated. The primary amide group forms hydrogen bonds in the S_1 subsite. The amide carbonyl of the inhibitor forms hydrogen bonds to NE2 of His-161 and the side-chain hydroxyl of Thr-142 and the amide NH_2 group forms a hydrogen bond with the backbone carbonyl of Thr-142. The S_1 subsite in RVP, unlike members of the chymotrypsin family, is not a well-defined pocket but rather

solvent exposed. The thiananaphthene ring makes VDW contacts with the S_2 subsite including a face to face interaction with the side chain of His-40.

E. Summary and Perspectives

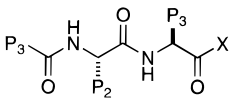
1. Reversible Inhibitors

While at first glance it appears the mechanism of peptide hydrolysis catalyzed by serine and cysteine proteases is quite similar, careful analysis of available data suggests that they are significantly different. While serine proteases have evolved to stabilize a highly charged transition state through oxyanion hole interactions, the evidence suggests that these interactions are less important for cysteine proteases.[322,323] Kinetic evidence suggest that for papain breakdown of the tetrahedral intermediate into an acyl enzyme, thioester formation is the rate-limiting step.[264] This is in contrast to serine proteases where both formation and breakdown of the tetrahedral intermediate are rate limiting.[134] On the basis of these mechanistic considerations, it would be expected that the types inhibitors which take advantage of the enzyme's ability to stabilize a transition state would be different for serine and cysteine protease inhibitors. That is, a class of inhibitors which are potent serine protease inhibitors, would not necessarily be potent cysteine protease inhibitors, and vice versa. This is in fact the case. Peptidyl nitriles, which by NMR experiments have been shown to resemble an acyl enzyme complex, are potent inhibitors of cysteine proteases but are feeble inhibitors, or substrates, for serine proteases. Peptidyl trifluoromethyl ketones, which by X-ray and NMR structural studies have been shown to resemble a charged tetrahedral intermediate are extremely potent inhibitors of serine proteases, but poor inhibitors of cysteine proteases. The only class of reversible inhibitors which are potent inhibitors of both serine and cysteine proteases are peptide aldehydes. Peptide aldehydes, which form diastereomeric covalent tetrahedral adducts with both serine and cysteine proteases, do not appear to be true transition state analogs, but rather "dead-end" inhibitors. A dead-end inhibitor refers to a compound that is converted by the enzyme into a stable adduct which is unable to proceed to a product because it lacks the requisite functionality for further processing.[152]

The most important lesson to be learned from the comparison of serine and cysteine protease inhibitors is that, despite apparent similarities, these two classes of enzymes are mechanistically distinct. Strategies which attempt to design inhibitors which take advantage of a cysteine protease's ability to bind a high-energy transition state must take into consideration the mechanistic differences between serine and cysteine proteases.

2. Irreversible Inhibitors: Quiescent Affinity Label Concept

The most potent small molecule inhibitors of cysteine proteases are affinity labels, such as α-chloromethyl ketones (**CP-25**) which irreversibly alkylate the active site cysteine. However, α-chloromethyl ketones are chemically reactive with simple thiols, such as glutathione, and various serine proteases,



CP-25 X = CH_2Cl

CP-26 X = CHN_2

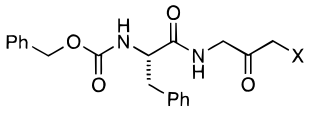
CP-27 X = CH_2F

CP-28 X = $CH_2S(CH_3)_2$

CP-29 X = CH_2OCOAr

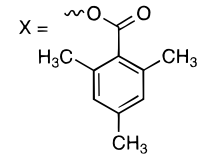
CP-30 X = CH_2OAr

Figure CP*19. Structures of irreversible cysteine protease inhibitors.

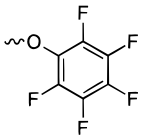


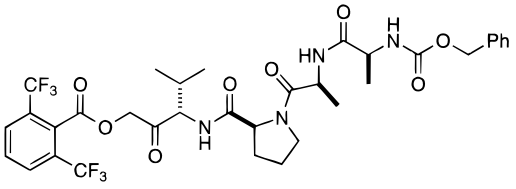
CP-31 X =Cl

CP-32



CP-33





CP-34

Figure CP*20. Structures of some quiescent affinity labels of cysteine proteases.

and are not selective for their target cysteine proteases. A potential drug candidate must be selective for its target enzyme and must be chemically stable to standard bionucleophiles such as glutathione. An ideal irreversible inhibitor is one which is uniquely reactive toward the active-site nucleophile of the target enzyme and quiescent in the presence of other bionucleophiles. These types of selective cysteine protease inhibitors are referred to as quiescent affinity labels.[271] Early examples of less reactive, more selective, affinity labels include diazo ketones **CP-26**,[260] α-fluoromethyl ketones **CP-27**,[283] and α-ketomethylsulfonium salts **CP-28**.[284] In a seminal publication, the Syntex group reported that α-[(arylacyl)oxy]methyl ketones **CP-29** and α-(aryloxy)-methyl ketones **CP-30** are selective, quiescent affinity labels of cysteine proteases of the papain class.[271] (See Figure CP*19.)

NMR studies have established that **CP-31**, **CP-32**, and **CP-33** (Figure CP*20) form an irreversible covalent adduct with papain. These studies indicate that all three inhibitors yield the same inhibited complex.[324] Since X-ray studies had previously shown that α-chloro ketones inhibit papain by alkylating the Cys-25 thiolate nucleophile, this study shows that the less chemically reactive inhibitors **CP-32** and **CP-33** also alkylate the thiolate by loss of mesitoic

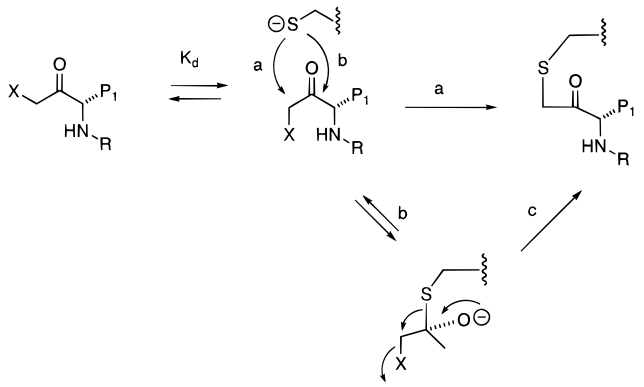


Figure CP*21. Mechanistic possibilities for reaction of the cysteine nucleophile with an affinity label reagent. Initial formation of a non covalent complex (K_d) can be proceeded by either (a) direct displacement of the leaving group X by the thiolate nucleophile or (b) reversible formation of a tetrahedral adduct which then (c) assists in the displacement of the leaving group X in the irreversible step.

acid or pentafluorophenol. The related compound **CP-34**, designed to be complimentary toward the serine protease HNE, was found to be a feeble inhibitor of this enzyme. Thus, compounds **CP-32** and **CP-33**, which are unreactive toward other bionucleophiles were demonstrated to be quiescent affinity labels of papain and cathepsin B.

As shown in Figure CP*21 there are two possible mechanisms for alkylation of the nucleophilic thiolate. The first mechanism consists of reversible formation of a noncovalent enzyme inhibitor complex (K_d) followed by direct nucleophilic displacement of the leaving group (a). The second mechanism involves reversible nucleophilic attack (b) by the thiolate on the ketone carbonyl, followed by migration of the sulfur (possibly by a distinct episulfonium ion intermediate) to displace the leaving group (c). Thus, speculation centers around whether a tetrahedral intermediate lies on the pathway to alkylated enzyme or whether a direct S_N2 displacement occurs. To date, there is no firm evidence to support, or rule out, either mechanism.[325] Both mechanisms require the initial formation of a noncovalent enzyme inhibitor complex and inhibitor efficiency will be a function of K_d. The success of the quiescent affinity label strategy relies on the ability of the enzyme to noncovalently bind the inhibitor in a manner that allows the powerful thiolate nucleophile to displace a poor leaving group. Efficient, irreversible inhibition of cysteine proteases by other weak electrophiles, such as epoxides and Michael acceptors, similarly require that the enzyme efficiently bind the inhibitor in a productive manner prior to irreversible chemistry.

The Sterling Winthrop group has reported studies that have extended the work on quiescent affinity labels. These studies have identified other leaving groups and have disclosed inhibitors which are selective for the cysteine protease ICE.[326−328] The Syntex group reported that **CP-35** is an efficient irreversible inhibitor of cathepsin B ($k/K = 1.7 \times 10^5$ M^{-1} s^{-1}). This compound was also found to be an effective in vivo cathepsin B inhibitor in rats. This result demonstrates that peptidyl (acyloxy)methyl ketones have promise as tools for the characterization of in vivo biochemical processes and as therapeutic agents.[265]

VI. Inhibition of Zinc Metalloproteinases

A. Introduction

One of the major classes of proteinases, zinc metalloproteinases are comprised of a number of families with varying degrees of structural similarity between them.[329−334] From a structural as well as mechanistic viewpoint, carboxypeptidase A (CPA) and thermolysin (TLN) have been the most extensively studied zinc metalloproteinases but are not in themselves targets for drug design. Furthermore, the structural basis of inhibition of both CPA[335,336] and TLN[337] have been reviewed, and further mention of these two enzymes will be made only in comparative instances. From a pharmaceutical perspective, undoubtedly the best known zinc metalloproteinase is angiotensin-converting enzyme (ACE), whose inhibition represents one of the most clinically effective means to lower blood pressure. Although SAR for myriad ACE inhibitors have been reported in the literature,[338,339] the absence of direct structural information precludes detailed discussion of this enzyme. Another less prominent therapeutic target in the cardiovascular area, endothelin-converting enzyme (ECE),[340−342] is also not amenable to structural studies. Therefore, the focus of this section shall be on matrix metalloproteinases (MMPs), a family of zinc metalloproteinases whose importance as a therapeutic target continues to grow, and for whom extensive SAR and structural information has recently become available.

B. Matrix Metalloproteinases: Mechanism of Action

Matrix metalloproteinases are involved in the degradation and remodeling of connective tissues, and, as a family, exhibit proteolytic activity toward virtually all of the constituents of the extracellular matrix.[343,344] The activity of MMPs is normally tightly regulated through (1) transcriptional regulation by cytokines and growth factors, (2) secretion as inactive proenzymes which require specific activation, and (3) the presence of endogenous tissue inhibitors of matrix metalloproteinases (TIMPs). Misregulation of MMPs, not surprisingly, is believed to be a major factor in number of disease states characterized by unwanted degradation of connective tissue, including arthritis and tumor invasion and metastasis. Fourteen metalloenzymes have been well-characterized as MMPs in humans, including three collagenases, three stromelysins, two gelatinases, matrilysin, metalloelastase, and four membrane-type MMP isozymes (Table MP*1). The members of the family differ substantially in domain structure, but all five family members whose crystal structures (of the catalytic domains) have been solved possess the same core structure in their catalytic domains. Important differences in the surface loops, however, lead to significant difference in the specificity subsites, especially the S_1' subsite, as will be discussed in detail below.

Table MP*1. The Matrix Metalloproteinase Family

MMP number	enzyme name(s)	therapeutic areas
MMP-1	fibroblast collagenase (HFC), collagenase-1	arthritis(?), cancer (?) periodontal disease
MMP-2	gelatinase A (Gel A), 72 kD gelatinase, type IV collagenase, HFG	cancer, MS, stroke angiogenesis
MMP-3	stromelysin-1 (HSln)	cancer, arthritis
MMP-7	matrilysin	cancer
MMP-8	neutrophil collagenase (HNC) collagenase-2	
MMP-9	gelatinase B (Gel B), 92 kD gelatinase HNG	cancer, MS, stroke
MMP-10	stromelysin-2	
MMP-11	stromelysin-3	cancer
MMP-12	metalloelastase	emphysema
MMP-13	collagenase-3	arthritis
MP-14−17	membrane-type MMPs	cancer

Table MP*2. Variable Active-Site Residues in MMPs

residue no.	MMP number 1	2	3	7	8	9	13	role
178	G	D	G	G	N	D	S	side chain in S_3'
179	G	G	N	N	G	G	G	C═O H bonds P_3' NH
180	N	L	V	T	I	L	L	NH H bonds to P_1' C═O; side chain in α
210	Y	Y	T	I	Y	Y	Y	side chain in S_3'
214	R	L	L	Y	L	L	L	in S_1'
215	V	V	V	A	V	V	V	in S_1'
235	L	L	L	V	L	L	L	in S_1'
237	Y	A	Y	Y	Y	Y	F	C═O in S_1'
239	S	I	L	T	N	M		side chain in S_2'
241	T	T	H	G	T	R	T	S_1'

Residue numbering in the MMPs is complicated by insertions and deletions in the sequences of various family members. For the purposes of clarity and consistency, we shall define a common numbering system for the MMPs, nominally based on a MMP-1 numbering system used by some authors,[345−349] in which structurally aligned residues will possess the same number in all the MMPs. There are several residues in the active site that are absolutely conserved among all (human) MMPs: (1) the catalytic zinc is coordinated by His-218, His-222, and His-228; (2) Glu-219 functions as the general acid−base at the catalytic site; (3) the carbonyl oxygen of Ala-182 H bonds to the NH of the scissile amide bond; (4) Leu-181 and Tyr-240, whose side chains together form a "wall" dividing the S_1' subsite and the S_3' cleft, and who each contribute a donor NH for H bonding to the substrate; (5) Pro-238 and Ala-234, whose carbonyl groups project into the S_1' subsite. Other active-site residues which vary in the MMPs are listed in Table MP*2.

The remarkable efficiency of the MMPs as proteinases (k_{cat}/K_m values for heptapeptides of ca. 10^5 M^{-1} s^{-1})[350] has important implications with respect to the mechanism of proteolytic cleavage by zinc metalloproteinases, in that MMPs clearly possess the minimal catalytic apparatus required to carry out the hydrolysis of peptide bonds. The MMP catalytic site possesses only a zinc ion coordinated by three imidazoles, a glutamic acid side chain suitably oriented to serve as an acid−base catalyst, and a main-chain carbonyl that appears to be a critical H-bond acceptor. Importantly, the MMPs lack the additional H-bond donor functionality found in TLN and in CPA that had previously been thought to play an essential role in catalysis through stabilization of the incipient oxyanion.[334,335,337,351] The observation of a conserved water molecule in several inhibitor: MMP-7 complexes in a position to H bond to the incipient oxyanion has led to the proposal that the MMPs employ this structured water molecule to replace the electrophilic catalysis groups of TLN and CPA.[348] Furthermore, in both TLN and CPA, the catalytic zinc is coordinated to two imidazole ligands and one carboxylate ligand, and it has been proposed that the charged carboxylate ligands decrease the intrinsic Lewis acidity of the zinc ions in TLN and CPA relative to the MMPs.[349]

Currently accepted mechanisms for the proteolytic action of TLN and the MMPs are shown schematically in Figure MP*1.[352] Prior to substrate binding, the enzyme presumably binds water as the fourth zinc ligand, with a near tetrahedral geometry of the four ligands about the zinc. A crystal structure of apothermolysin (PDB entry:1LNF)[353] illustrates the likely binding mode of water in an apo-MMP, although it is interesting to note that in the structure of "apo-MMP-1" reported by the Glaxo group[352] as well as in the structure of "apo-MMP-3" solved by the Agouron group,[354] the zinc coordination sphere is filled not by water but by functionality from residues from adjacent protein molecules in the crystal lattice. Although the initial apoenzyme may be represented as possessing water bound to zinc, with strong H bonding to the carboxylate of Glu-219, it is equally plausible that proton transfer from the zinc-bound water to the carboxylate has occurred to give a zinc hydroxide species accepting a strong H bond from the carboxylic acid function of Glu-219. In either case, however, it is clear that the charge of the Zn^{+2} ion is not fully compensated in the catalytic site of the MMPs, in contrast to TLN and CPA, where the presence of a carboxylate ligand on zinc leads to an overall charge neutral arrangement. Coordination of the carbonyl oxygen of the scissile peptide bond to the zinc is followed by nucleophilic attack of the zinc-bound water with simultaneous (or prior) proton transfer to Glu-219 to form the zinc-bound tetrahedral intermediate. Subsequent proton transfer to the scissile nitrogen from the carboxyl oxygen of Glu-219 which is oriented toward the "primed" binding sites (hereafter referred to as Oϵ1) appears facile. Upon breakdown of the tetrahedral intermediate, a second proton transfer from the hydroxy group to

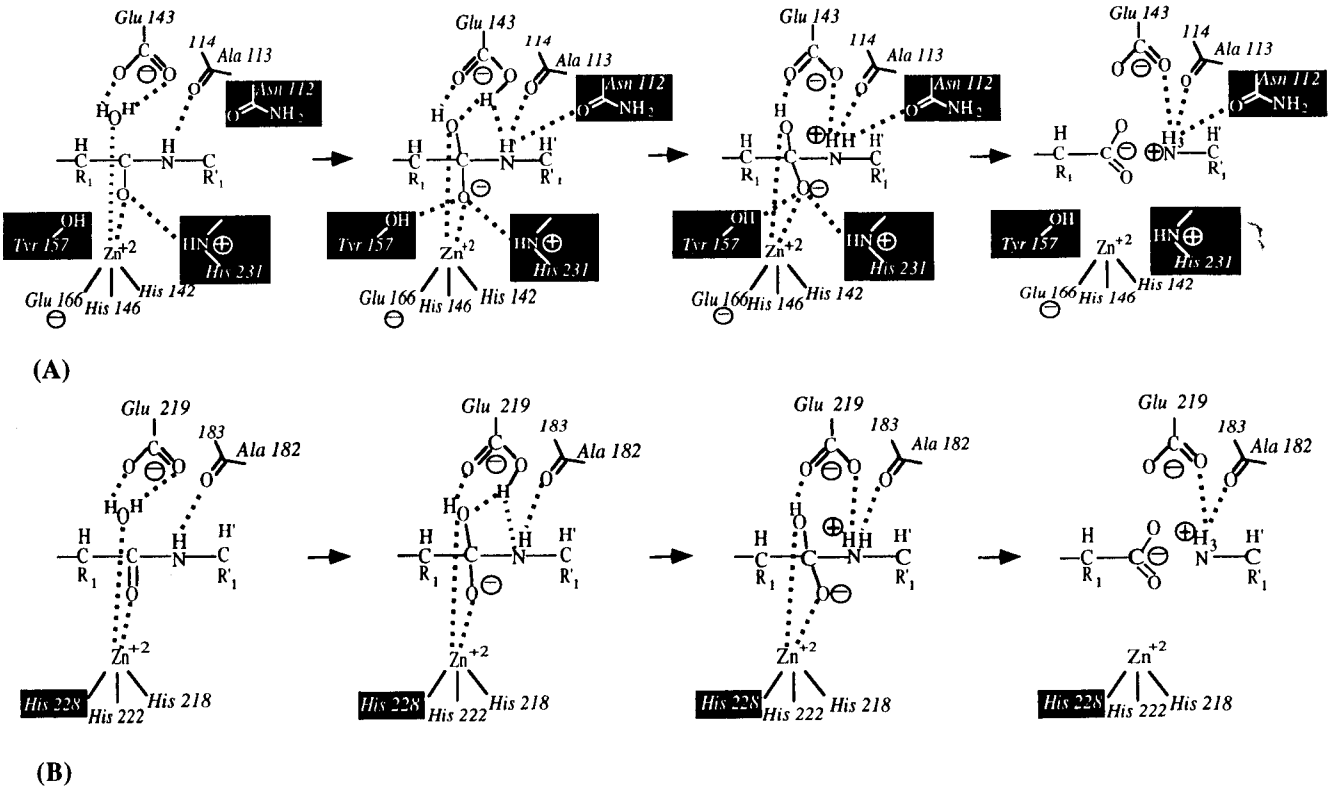


Figure MP*1. Reaction mechanisms for the proteolytic action of thermolysin (A) and MMP-1 (B). Functional groups unique to either thermolysin or MMP-1 are highlighted. (Reproduced from ref 352. Copyright 1994 American Chemical Society.)

Glu-219 yields the final complex of the carboxylate bound to zinc with a H bond to protonated Glu-219 and the neutral amine, or alternatively, as shown in Figure MP*1, an additional proton transfer to the leaving amine generates the ammonium ion salt bridged to Glu-219. It has been generally observed that the predominant binding determinants of peptide substrates for zinc metalloproteinases reside on the "prime" side of the scissile bond, and that "unprimed" side binding is generally weak. This predilection is most likely due to the need to avoid strong product inhibition; if "unprimed" binding was tight, then the carboxylate product, with an effective zinc ligand, would tend to be a good inhibitor. Obviously, this principle dictates inhibitor design, as we shall discuss in the following paragraphs.

C. Inhibition of Matrix Metalloproteinases

1. Classes of MMP Inhibitors

All but one of the MMP inhibitors considered in this review possess a zinc-binding function (ZBF) attached to a framework that binds to the "primed" binding regions. These inhibitors have been referred to right-hand-side (RHS) inhibitors in reference to the convention of drawing the prime residues of a peptide substrate on the right side.[355] The vast majority of inhibitors fall into two classes of peptidic inhibitors: (1) Class I inhibitors, which include succinamide carboxylates and hydroxamates, possess two sp^3 carbons between the ZBF and the first peptide bond; (2) class II inhibitors, which include carboxyalkyl α-amino amides (Z = NH) and glutaramide carboxylates (Z = CH_2), possess three sp^3-hybridized atoms between the ZBF and the first peptide bond. As shown in Figure MP*2, class I inhibitors are conceptually derived by replacing the scissile peptide bond with a ZBF-CH_2 group, such that the ZBF is geometrically positioned to coordinate to zinc in analogy to the oxygens of the primary tetrahedral intermediate **A**. Class II carboxyalkyl α-amino amide inhibitors, on the other hand, have been regarded as "collected-product analogs" but also may be described as mimicking the interactions of the secondary tetrahedral intermediate **B** (Figure MP*2). The third class of MMP inhibitors whose binding mode has been structurally characterized is comprised of *N*-(arylsulfonyl)-α-amino *N*-hydroxy amides and *N*-hydroxy malondiamides, which both bind to MMPs in a similar manner that is not substrate like. Despite the rarity of LHS inhibitors, crystallographic information is available for one such inhibitor, Pro-Leu-Gly-NHOH, and it shall be briefly discussed near the end of this section. Several reviews of MMP inhibitors have appeared recently which cover SAR trends in detail,[344,355−359] but the structural basis for inhibition has not been previously reviewed.

2. Class I Pseudopeptide Inhibitors

a. Overview of SAR. For class I inhibitors, many different ZBFs have been explored, and the hydroxamate function provides the best potency by far. Roughly 5−10-fold less potent than hydroxamates are "reverse hydroxamates", while thiols are typically 20−50-fold less potent and carboxylates are 100−2000-fold less potent.[360] Phosphonic acids appear to be comparable in potency to carboxylates,[361] and phosphinates are reported to be less potent still, although additional binding may be obtained through

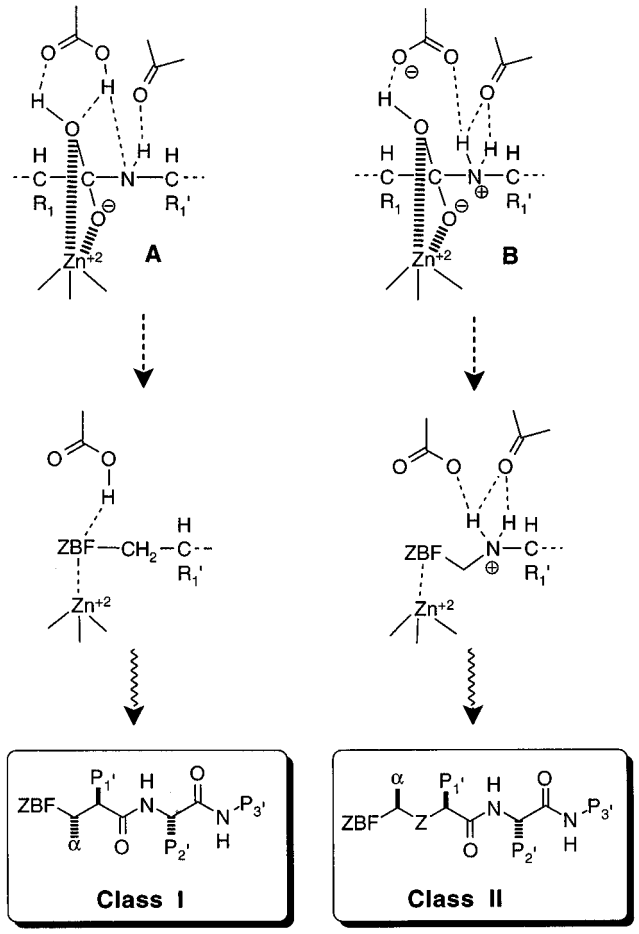


Figure MP*2. Design features of class I and class II inhibitors relative to primary (**A**) and secondary (**B**) tetrahedral intermediates.

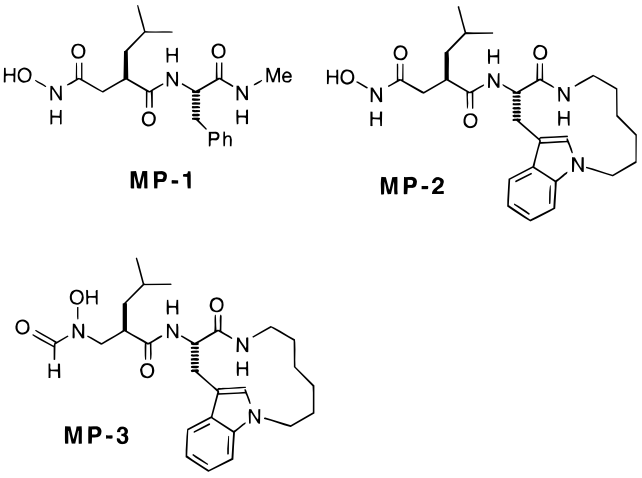


Figure MP*3.

incorporation of certain large groups on the unprimed side.[360,362,363] The least potent among ZBFs that have been structurally characterized is the sulfodiimine group.[348] A number of other potential ZBFs, including amidoxime, 2-pyridylamide, and 2-imidazolylamide, are reported to not show useful levels of binding.[361]

For class I dipeptide hydroxamates, inhibitors with an isobutyl group at P_1', a hydrophobic side chain at P_2', and a simple *N*-methyl amide "cap" at P_3', as exemplified by **MP-1** (Figure MP*3), potently inhibit all of the MMPs with K_i's in the 1−100 nM range, with MMP-3 generally being the least well inhibited. Replacement of isobutyl with certain large substituents at P_1',[364,365] not corresponding to natural amino acid side chains, lead to more potent inhibitors for most of the MMPs, with the exception of MMP-1 and MMP-7, wherein activity is diminished. In general, relatively small effects on potency are observed in substitution at P_2' and P_3', although particular substituents have been found to provide useful levels of selectivity for certain enzymes.

b. Interactions of Class I Inhibitors at the Zinc Site. The zinc binding modes of the numerous class I hydroxamates for which structural data are available are remarkably consistent; the crystal structure of the prototypical inhibitor **MP-1** bound to MMP-1 was determined at 1.56 Å resolution (PDB entry: 1HFC) and serves well to illustrate the key interactions with all of the enzyme functionality involved in catalysis (Figure MP*4). In solution, the hydroxamic acid group is an ambident acid, with similar pK_a's for the terminal OH and the NH which are generally in the range of 8−10. ^{15}N NMR studies of the binding of an ^{15}N-labeled hydroxamate to MMP-3 has established that the hydroxamate binds as an deprotonated O-acid.[366] It is worth noting that Holmes and Matthews concluded, on the basis of crystallographic analysis, that a class III hydroxamic acid inhibitor was bound to TLN as its N-deprotonated anion, without a H bond to Ala-113 (TLN numbering).[367]

As shown in Figure MP*4, the hydroxamate binds with bidentate ligation to zinc to form a slightly distorted trigonal-bipyramidal coordination geometry at zinc. The hydroxamate oxyanion forms a strong, short H bond (2.5 Å in 1HFC) to the carboxylate oxygen of Glu-213 that is oriented toward the unprimed binding regions (hereafter referred to as Oϵ2). A H bond between the hydroxamate NH and the carbonyl oxygen of Ala-182 also contributes to binding. Thus, a set of strong interactions is achieved at the zinc site, without any significant unfavorable contacts. Table MP*3 summarizes geometrical values for polar interactions of class I inhibitors with the active site of MMPs.

An alternative hydroxamate connectivity, wherein the nitrogen is bonded to the inhibitor backbone and commonly referred to as "reverse hydroxamate", is present in **MP-3**. The crystal structures of **MP-3**:MMP-7 complex and the corresponding "normal" hydroxamate **MP-2** bound to MMP-7 (PDB entry: 1MMQ) have been determined and, not surprisingly, the structural differences are small and comparable to the experimental error at this level of resolution. Nevertheless, the differences are in the expected direction, with a slight opening of the catalytic zinc site so as to lessen the close contact between carbonyl carbon of the reverse hydroxamate **MP-3** and the carbonyl oxygen of Ala-182 (3.08 Å vs H−N to O=C H-bond distance of 2.91 Å for **MP-2**).

Although the hydroxamate function appears to be the optimal zinc ligand for the MMPs, concerns about the unfavorable pharmacokinetics (poor oral bioavailability, rapid clearance) of the early class I hydroxamates, as well as continuing concerns over the potential for chronic toxicities arising from metabolic activation of the hydroxamate group, have

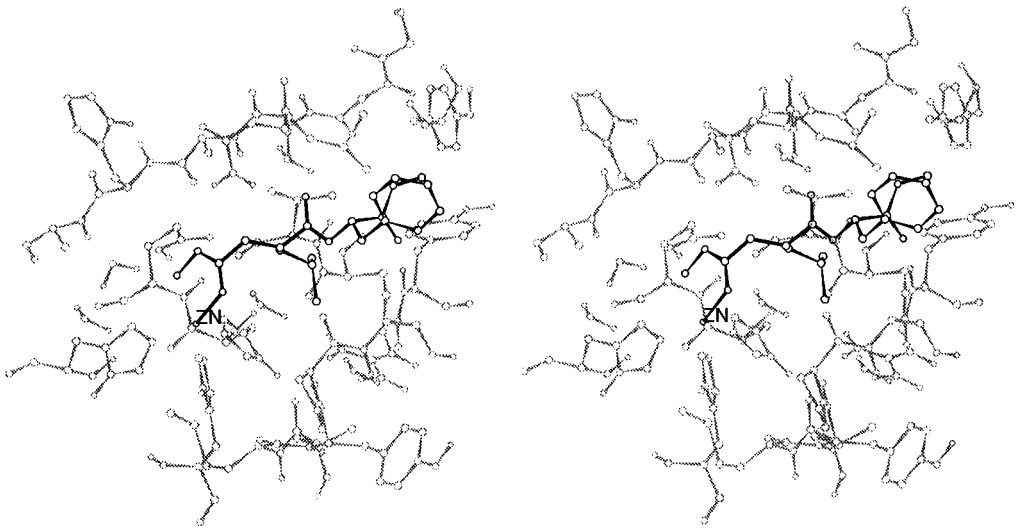
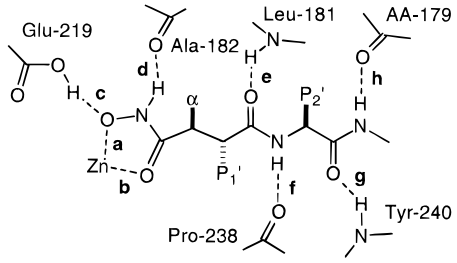


Figure MP*4. Stereoview of the prototypical class I dipeptide hydroxamate inhibitor **MP-1** (black) bound to MMP-1 (gray).

Table MP*3. Zinc−Ligand and Hydrogen Bond Distances for MMP:Class I Inhibitor Complexes



					distance (Å)[a]							
entry	PDB file	enzyme	resolution	K_i (nM)	a	b	c	d	e	f	g	h
MP-1	1HFC	MMP-1	1.56	7	2.2	2.1	2.5	2.8	2.8	3.1	2.8	2.8
MP-1	1MNC	MMP-8	2.1	2	2.3	2.1	2.4	3.0	2.9	2.9	2.9	2.8
MP-2	1MMQ	MMP-7	1.9	1	2.2	2.2	2.6	2.9	2.9	3.0	2.8	2.9
MP-3		MMP-7		18	2.2	2.1	2.65	NA[b]	2.8	3.0	2.8	2.8
MP-4	1MMP	MMP-7	2.4	570	2.4	2.1	2.8	NA	2.8	3.1	2.8	2.9
MP-5		MMP-1	2.1	7	2.7	2.0	2.9	NA	2.9	3.2	2.7	2.8
MP-6	1MMR	MMP-7	2.3	1300	2.8	2.0	2.95	NA	2.8	3.0	2.8	2.8
MP-7		MMP-7		1800[c]	3.3	2.0	2.8	NA	2.85	3.3	2.9	2.9
MP-8	1JAO	MMP-8	2.4	1200	NA	NA	NA	NA	2.8	3.1	2.9	2.9
MP-10		MMP-7	2.3	18	2.2	2.0	2.7	2.8	3.1	2.8	2.9	2.9
MP-11		MMP-7	1.9	5000	2.2	2.0	3.0	NA	2.8	3.2	2.7	3.1
MP-12		MMP-3	2.0	1.5	2.1	2.2	2.8	3.0	2.6	3.3	2.9	NA
MP-13		MMP-3	2.0	170[d]	2.3	2.1	2.9	2.8	2.7	NA	NA	NA
MP-14		MMP-3		0.6	2.1	2.1	2.8	3.0	2.9	2.9	2.9	2.8
MP-15		MMP-7	2.1	4900	2.4	2.0	2.8	NA	2.8	3.2	2.8	3.0
MP-16		MMP-3	2.5	1.9	2.3	2.0	2.9	NA	2.9	3.0	2.8	3.0
MP-16		MMP-7		610	2.5	2.0	2.6	NA	2.8	3.3	2.9	2.8
MP-17		MMP-3	2.3	14	2.6	2.0	2.7	NA	2.9	2.8	2.7	2.9
MP-18		MMP-3	2.2	2.7	2.5	2.0	2.7	NA	2.9	2.9	2.8	2.8
MP-20		MMP-1	1.9	1900	2.9	2.15	2.6	NA	2.8	3.1	2.9	2.8
MP-24	1TLC	MMP-1	2.2	(9)	2.4	2.1	2.7	3.0	3.0	3.1	2.8	2.8
MP-26	1MMB	MMP-8	2.1	0.6	2.7	2.0	2.4	3.0	2.8	3.3	3.0	3.1
MP-26		MMP-1	2.5	0.2	2.15	1.9	2.7	2.8	2.9	2.9	2.7	3.2

[a] Heavy atom to heavy atom distances (e.g., Zn−O or N−O). [b] NA indicates participating group on the inhibitor is not present. [c] IC_{50} for MMP-1. [d] 1:1 mixture of diastereomers.

instigated extensive investigation into alternative zinc ligands. First, we shall consider the carboxylate group, since the most structural information, after hydroxamates, is available for this class of zinc ligand. Class I carboxylate inhibitors are generally 100−2000-fold less potent (at pH 7.5) than the corresponding hydroxamates,[361,364] with an interesting general trend that the factor tends to decrease as the potency of the carboxylate inhibitor improves. In the inhibition of MMP-3, for example, the factor for carboxylate−hydroxamate inhibitor pairs when the P_1' group is isobutyl is 1000−2000, but when the P_1' group is biphenylpropyl, which improves potency about 1000-fold for the carboxylate inhibitor, the carboxylate−hydroxamate potency ratio drops to ∼100.[368]

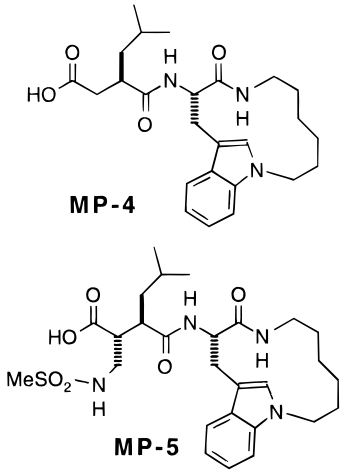


Figure MP*5.

In view of this SAR, it is not surprising that class I carboxylates bind in a manner that positions the inhibitor backbone in a very similar, although not identical, manner to that observed with the corresponding hydroxamates. The carboxylate−hydroxamate pair **MP-4** and **MP-2** provides a clear illustration: the carboxylate group of **MP-4** binds to MMP-7 (PDB entry: 1MMP; Figure MP*5) in an unsymmetrical, bidentate fashion to the zinc, with one oxygen at the position occupied by the carbonyl oxygen of the hydroxamate, 2.08 Å from the zinc.[348] The other carboxylate oxygen is 2.40 Å from the zinc, which is about 0.65 Å closer to the zinc than the hydroxamate nitrogen to which it corresponds structurally. This shift results in a ca. 0.5 Å movement of the backbone of the carboxylate inhibitor toward the zinc atom, relative to the hydroxamate inhibitor. Because the backbone is more or less extended along this direction and the backbone H bonding occurs in a β-sheet-like manner (*vide infra*), the key backbone H bonds are maintained.

It appears that several factors contribute to the 3−5 kcal/mol advantage in hydroxamate binding over carboxylate binding; the most important factor appears to be the difference in acidity. Perhaps contrary to intuition, carboxylate inhibitors bind more tightly, about 1 log unit per pH unit, to MMPs as the pH is lowered, whereas hydroxamate binding is essentially pH independent (pH range 5−8). This is readily rationalized by considering the overall binding process wherein the zinc-bound water is released upon inhibitor binding, leaving Glu-219, in a formal sense, unprotonated (Figure MP*6). Upon binding of the neutral hydroxamic acid form of the inhibitor, proton transfer from the terminal OH to the Oϵ2 oxygen (*vide supra*) leads directly to the observed bound state **III-h**. On the other hand, binding of the anionic carboxylate inhibitor to give, initially, **II-c** must be followed by the binding of a proton (from bulk solution) to the Oϵ1 carboxylate oxygen of Glu-219 in order for the bound state **III-c** to be achieved. Another way of looking at this, from a thermodynamic view, is that the neutral carboxylic acid form is binding to the enzyme (i.e., **I** ⇒ **III-c** in Figure MP*6). Thus, the 2−3 log unit difference in binding affinity is similar to the 3 pH units between the pK_a of the carboxylates and the assay pH of 7.5, and the intrinsic affinity of hydroxamic acid inhibitors and carboxylic acid inhibitors may well be similar. In practical terms, however, carboxylate inhibitors are much weaker than hydroxamate inhibitors under physiological conditions, near neutral pH.

Another factor that may play a role in the diminished potency of class I carboxylate inhibitors is the loss of the H bond of the hydroxamic N−H to the carbonyl oxygen of Ala-182. Interestingly, the cocrystal structures of all of the class I carboxylate inhibitors in Chart I except **MP-5** show an ordered water molecule at H-bonding distance to the carbonyl oxygen of Ala-182. Compound **MP-5** was designed to position the N−H of the sulfonamide group in the α-side chain so as to replace the bound water, and the structure of **MP-5** bound to MMP-1 (Figure MP*7) confirms that the H bond is present (N···O=C

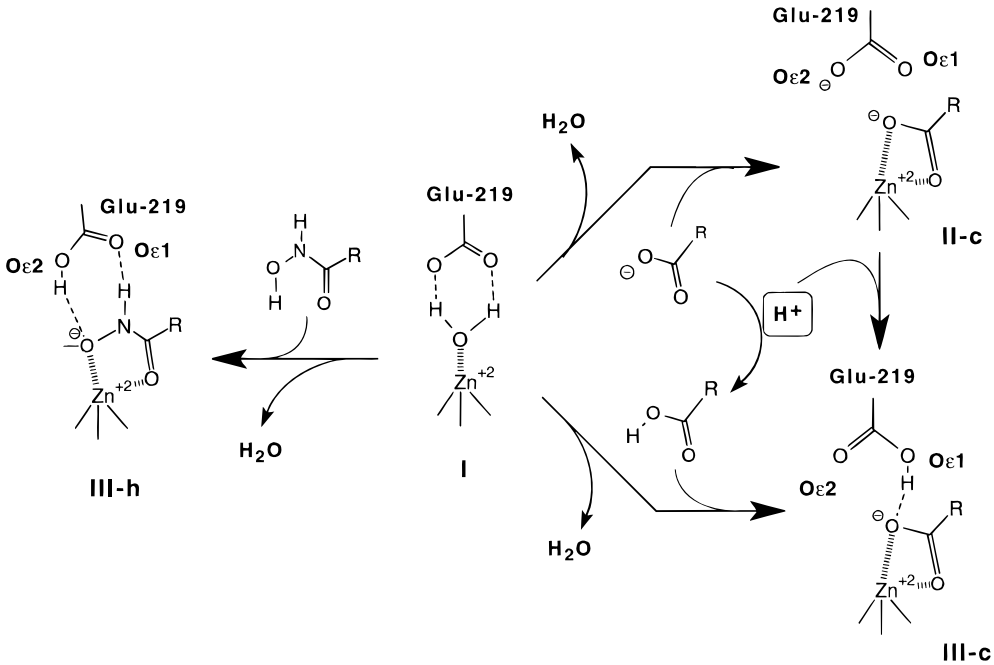


Figure MP*6. Protonation states of hydroxamates and carboxylates during binding.

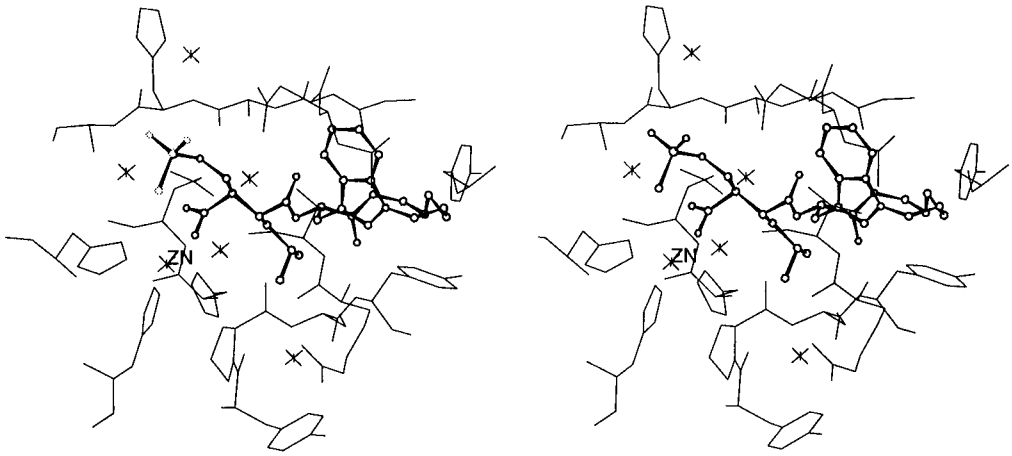


Figure MP*7. Stereoview of **MP-5** (ball and stick) bound to MMP-1.

distance of 3.0 Å).[369] Contrary to expectations, **MP-5** is only modestly more potent (ca. 2−10-fold for several of the MMPs) than **MP-4**; it is, of course, not possible to delineate the precise energetic contribution of the H bond relative to the effects of the substitution on other factors such as van der Waals interactions with the protein, and inhibitor desolvation and preorganization.

Within the class I inhibitor group, three other ZBFs—phosphinate, sulfodiimine, and thiolate—have been determined by X-ray crystallography to bind in a monodentate fashion to zinc. In the complex of phosphinate **MP-6** with MMP-7,[370] the "outside" (toward solvent) oxygen of the phosphinate is coordinated to zinc at a distance of 1.96 Å, while the "inside" oxygen is 3.32 Å from zinc but is clearly H bonding (2.82 Å) to the carboxylate Oϵ1 oxygen of Glu-219. This ligation geometry is essentially the same as previously determined for the binding of phosphorus-based inhibitors to TLN[337] and CPA.[371,372] Because of the longer C−P and P−O bond lengths, the CH_2 group attached to phosphorus is displaced toward the carbonyl oxygen of Ala-182, resulting in a close contact distance of 3.18 Å, relative to the 3.4−3.5 Å distances observed in hydroxamate and carboxylate inhibitor complexes.

The sulfodiimine **MP-7** also binds in a monodentate manner with the "outside" imino nitrogen ligated (2.0 Å) to zinc and the "inner" imino nitrogen H bonding to Glu-219 (PDB entry: 1MMR; Figure MP*8).[348] Furthermore, superimposition of the **MP-7** complex and the **MP-6** complex reveals that the sulfur atom occupies very nearly the same location as the phosphorus atom (∼0.3 Å shift). Nevertheless, there are significant differences in the details of the binding interactions of the two ZBFs which seems to be driven by a difference in H bonding to Glu-219. In the sulfodiimine complex, the inner imino nitrogen is (presumably) protonated and donating a H bond to the Oϵ2 oxygen of Glu-219; the observed ONS angle of 132° is consistent with a sp^2-hybridized nitrogen and a reasonable H-bond geometry (N—H···O angle ∼130°). It is unclear whether Glu-219 is unprotonated or protonated on the Oϵ1 oxygen; studies of the pH dependence would presumably resolve this issue, but none have been reported. On the other hand, the "inner" oxygen of the phosphinate **MP-6** accepts a H bond from the protonated Oϵ1 oxygen of Glu-219 (O—H···O angle ∼173°) without particular constraints on the P═O···H angle. Thus, if the sulfodiimine were to be in an identical geometry as the phosphinate, the N−H would be unable to make a favorable H bond to the Oϵ1 oxygen of Glu-219 (N—H···O angle ∼85° based on modeling). It is worth noting that a sulfodiimine inhibitor has been reported to bind to CPA in an alternative fashion, in which the "inner" imino nitrogen is ligated to zinc (2.2 Å) and is presumably accepted a H bond from protonated Glu-270 (CPA numbering).[373] In this case, the "outer" imino nitrogen is 3.4 Å from zinc and is presumably protonated.

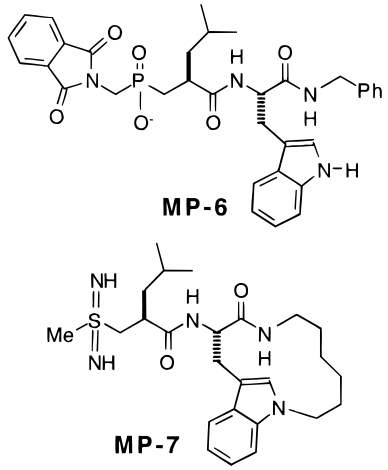
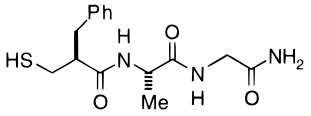


Figure MP*8.

Although the intrinsic affinity of the monodentate thiolate ZBF for zinc is probably less than that of oxygen-based bidentate ZBFs, the low desolvation cost as well as the ease of ionization are strongly compensating factors. In general, the potency of class I thiol inhibitors against MMPs is intermediate between that of hydroxamate- and carboxylate-based inhibitors. Structural details of thiolate binding are available from the complex of MMP-8 with the inhibitor **MP-8** (PDB entry: 1JAO; Figure MP*9).[374] In addition, the structure of prostromelysin (pro-MMP-3) has been reported (PDB entry: 1SLM),[375,376] in which Cys-75 of the propeptide provides a ligating thiol group. Despite marked differences in the overall way the propeptide binds to the active site relative to the inhibitor, the position of the sulfur atom bound to zinc is nearly identical (see Table

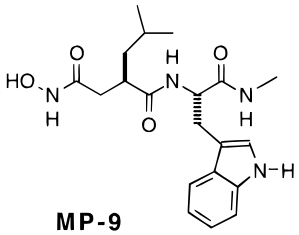


MP-8

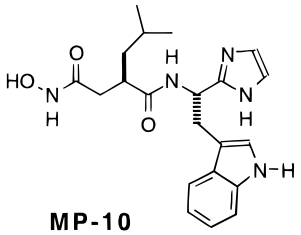
Figure MP*9.

Table MP*4. Comparison of Zinc−Thiolate Interaction Geometries

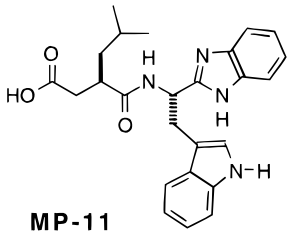
interaction	pro-MMP-3	MMP-8:**MP-8**
Zn−S (Å)	2.21	2.24
Glu(219)Oϵ1−S (Å)	3.17	3.03
His(218)Nϵ2−Zn−S (deg)	118.9	109.3
His(222)Nϵ2−Zn−S (deg)	114.5	109.4
His(228)Nϵ2−Zn−S (deg)	115.3	124.7
Zn−S−C (deg)	96.3	124.2
Zn−S−C−C (deg)	168.8	48.3



MP-9



MP-10



MP-11

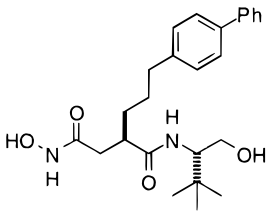
Figure MP*10.

MP*4). In both cases, a very slightly distorted tetrahedral geometry is observed, with a Zn−S bond distance of about 2.2 Å.

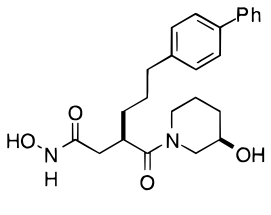
c. Backbone Interactions of Class I Inhibitors with MMPs. As illustrated in Table MP*3, the P_1'/P_2' and P_2'/P_3' amide groups of class I dipeptide inhibitors typically make four strong H bonds with the protein backbone, essentially filling in one strand of a kinked β-sheet-like structure. Of these interactions, the H bond between the P_2' N−H and the C=O of Pro-238 may be the weakest, since in certain structures the distance has been observed to lengthen to >3.3 Å.

Invariably, attempts to substitute for the amide bonds in class I inhibitors have led to a substantial loss in potency, although certain surrogates for the P_2'/P_3' amide have been found to be somewhat successful. Structural information is available for two of these surrogates: an imidazole ring (**MP-10** and **MP-11**, Figure MP*10), which was designed to make both H bonds,[347] and the primary alcohol group (**MP-12**, Figure MP*11), which makes only one H bond.

Substitution of the C-terminal *N*-methyl amide group of galardin (**MP-9**) with an imidazole gives **MP-10**, with a loss of potency of about 5-fold for all MMPs tested. Although the crystal structure of a **MP-9**:MMP complex is not available, comparison of the cocrystal structures of **MP-2** and **MP-10** with MMP-7 clearly shows that the imidazole substitutes



MP-12



MP-13

Figure MP*11.

for the amide, making H bonds to the carbonyl of Asn-179 (2.89 Å) and to the N−H of Tyr-240 (2.89 Å), without any detectable perturbation of the position of the inhibitor framework and without introducing any new close contacts.[347] On the other hand, the benzimidazole **MP-11** does not superimpose as well with the corresponding carboxylate-based inhibitor **MP-4**; the benzimidazole ring system of **MP-11** is less embedded in the S_3' cleft than is the amide group of **MP-4**, resulting in a longer (3.15 Å) hydrogen bond from the N−H of Tyr-240 to the imidazole imino nitrogen.[347] This appears to be a consequence of the close contact (of ca. 2.8 Å, according to modeling) that would result between C-4 of the benzimidazole and the carbonyl oxygen of Tyr-240 if the benzimidazole were to bind in the same position as the amide group. In the observed structure of **MP-11** bound to MMP-7, this distance is 3.2 Å.

The replacement of a C-terminal N-methyl amide with a hydroxyl group generally results in a 10−50-fold loss of activity, reaffirming the importance of the backbone interactions.[377] In principle, the C-terminal hydroxy group of the inhibitor **MP-12** (Figure MP*11) could either donate a H bond to carbonyl oxygen of Asn-179, or accept a H bond from the NH of Tyr-240. In the crystal structure with MMP-3, the hydroxy group acts as an acceptor; one possible explanation for this is a relatively close contact of ~3.1 Å that would result between the hydroxyl oxygen and one of the methyl groups of Leu-181 if the hydroxyl were to H bond to the carbonyl oxygen of Asn-179.

The more profound backbone modification in **MP-13**, in which the P_1'/P_2' amide is N,N-disubstituted, results in an even more substantial loss in binding: the K_i of **MP-13** for MMP-3 is 170 nM compared to 1.5 nM for the alcohol **MP-12**. In the structure of the **MP-13**:MMP-3 complex, the active-site cleft has opened up in the P_2'/P_3' area, such that the distance from the carbonyl oxygen of Pro-238 to the nitrogen of Leu-181 (normally the acceptor−donor pair for the P_1'/P_2' amide bond) is 8.6 Å compared to 7.8 Å in the cocrystal structure of **MP-12** and other class I inhibitors (see Figure MP*12). Even with this expansion, a close contact of 3.0 Å is present between the Pro-238 oxygen and the methylene group of **MP-13** that is replacing the normally H-bonding hydrogen atom. It is noteworthy that the position of the entire "succinate" unit of **MP-13**, including the hydroxamate group and the P_1' side chain, are virtually superimposable in the **MP-12** and **MP-13** complexes. In addition, the hydroxyl group on the C-terminal ring makes no polar interactions with the protein, but simply projects into solvent.

d. Binding Interactions in the S_1' Subsite. Although the ZBF is clearly the primary determinant

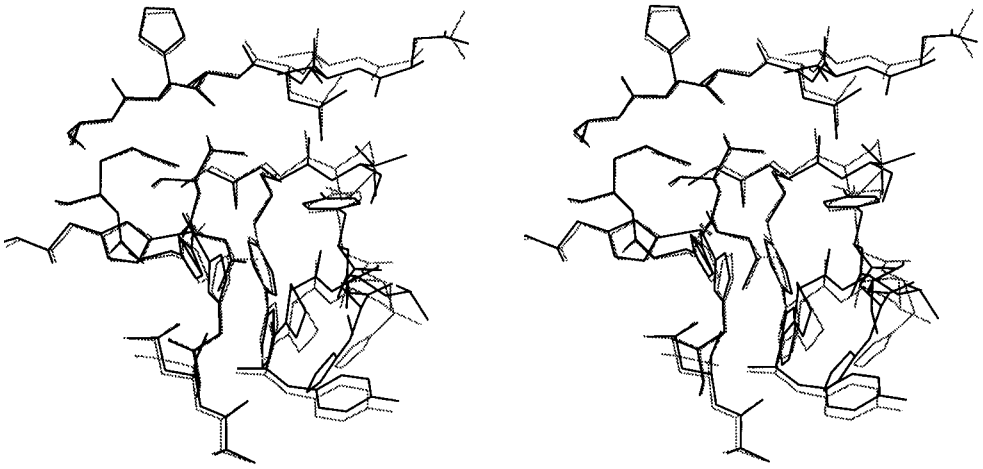


Figure MP*12. Stereoview of the superposition of **MP-12**:MMP-3 (black) and **MP-13**:MMP-3 (gray) structures.

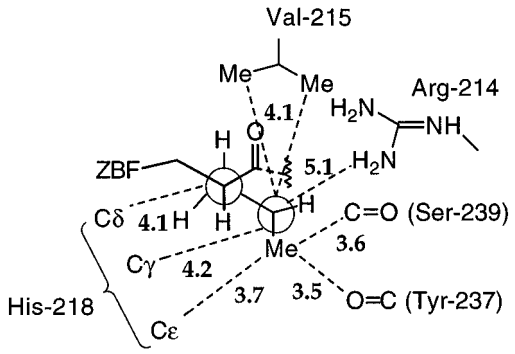


Figure MP*13. Newman projection of class I inhibitor isobutyl P_1' conformation with VDW contacts; distances are in Å.

of inhibitor potency in a relatively nonselective manner, the P_1' side chain has also been found to not only be critical for potency but also can dramatically influence the enzyme selectivity profile of an inhibitor. As an example of nonselective binding effects, replacement of the P_1' isobutyl group in **MP-9** with a methyl group results in a 4.2 kcal/mol decrease in binding affinity to MMP-3, with similar effects on other MMP enzymes.[378] On the other hand, replacing the P_1' isobutyl group of **MP-4** with biphenyl propyl leads to a 4.1 kcal/mol *increase* in binding to MMP-3, while resulting in a 1.7 kcal/mol *decrease* in binding to MMP-1.[368] As will be discussed below, the S_1' subsite of the MMPs is largely hydrophobic and highly variable in size and shape among the different members of the family.

Early work on class I inhibitors directed toward MMP-1 quickly settled on the isobutyl side chain as a highly effective P_1' group, in direct analogy to the preference of MMP-1 for Leu (and Ile) at the P_1' position of peptide and protein substrates,[379] and, as a result, the majority of MMP inhibitors prepared to date possess isobutyl as the P_1' side chain. In the last few years, several crystal structures of MMP-1 complexed to class I inhibitors bearing P_1' isobutyl groups have been solved which illustrate clearly that the S_1' subsite of MMP-1 is moderately sized, such that the isobutyl group fits snugly with several favorable VDW contacts, as shown schematically in Figure MP*13 for the **MP-1**:MMP-1 complex. Perhaps surprisingly, the "floor" of the pocket is comprised of the guanadinium group of Arg-214, which is not ion paired although it does make several H bonds, including one to an ordered water molecule. Although an edge of the guanadinium group is exposed to the pocket, all attempts thus far to design improved P_1' groups that H bond to the guanadinium group have apparently failed,[365] perhaps due to the inability to achieve optimal angles for the desired H bonds.

As additional members of the MMP family were discovered and characterized with respect to inhibition, it became evident that while all family members tolerated P_1' isobutyl in class I inhibitors, certain members of the MMP family preferred larger P_1' groups. For example, prior to elucidation of the structure of any MMP, the Celltech group discovered that the phenylpropyl group as the P_1' side chain enhanced potency for MMP-2 and MMP-3 while dramatically decreasing potency for MMP-1, presumably due to blockage by Arg-214.[361,364] Subsequently, determination of the crystal structures of MMP-3 and MMP-8 revealed that these enzymes possessed very large, predominately hydrophobic S_1' "tunnels" that passed all the way through the protein. By sequence alignment[380] as well as by consideration of the existing SAR, it may be concluded that all other MMPs with the exception of MMP-7 and possibly MMP-11 also possess large S_1' subsites in their ground state conformation. In the case of MMP-7, the side chain of Tyr-214 resides in a similar position as the guanadinium group of Arg-214 in MMP-1, such that the Tyr-214 phenolic OH is nearly superimposed on one of the terminal NH2 groups of Arg-214 when MMP-1 and MMP-7 (in complex with P_1'-isobutyl containing inhibitors) structures are overlaid.

One side of the large S_1' "tunnel" in the other MMPs consists of residues from the core of the protein, including the face of the zinc-ligating imidazole ring of His-218. The other side of the S_1' subsite is formed from an extended loop, and the loop is variable in both length and sequence composition among the various MMPs. Not surprisingly, the core side of the "tunnel" is essentially superimposable in crystal structures of MMPs complexed with different inhibitors, whereas the loop side appears to be somewhat flexible. Flexibility in this region has also been observed in two MMPs with "small" S_1' pockets, MMP-1 and MMP-7, and the important consequences

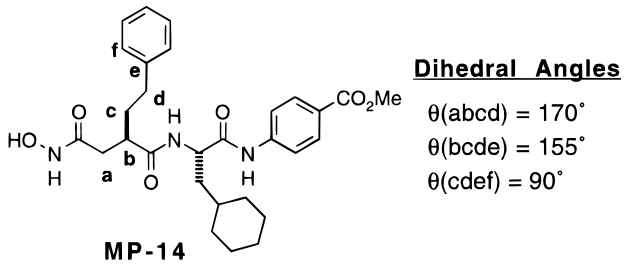


Figure MP*14. Structure of **MP-14** and dihedral angles of its bound conformation when complexed to MMP-3.

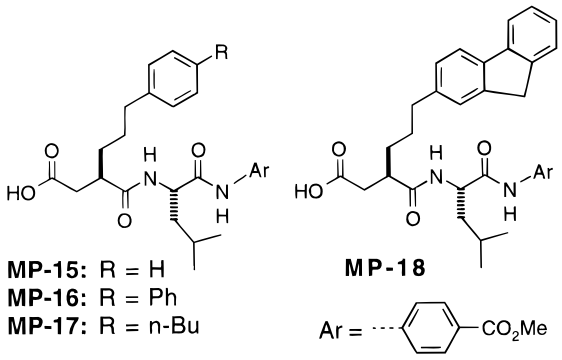


Figure MP*15.

of this conformational mobility will be discussed in detail below.

The X-ray structure of **MP-14** bound to MMP-3 shows that nonoptimal torsions in the ethyl linker, particularly the b−c−d−e dihedral of 155° (Figure MP*14), are needed in order to position the phenyl ring of phenethyl P_1' side chain in optimal VDW contact.[368] In this case, the potentially large gain in burying the phenyl ring in a hydrophobic environment is at least partially opposed by an increase in torsional strain. It is also clear that much of the S_1' subsite remains unfilled by the phenethyl group, but the bond vectors from the 3- and 4-positions on the phenyl ring do not appear to be well-suited to access this unfilled space. This analysis is supported by results from the Merck group with class I phosphinate inhibitors, who found that methoxy substitution at the 3- or 4-positions of the phenyl ring resulted in slight loss of potency against MMP-3.[362]

Although potencies of the carboxylic acid analog of **MP-14** (MMP-3 K_i = 340 nM) and **MP-15** (MMP-3 K_i = 310 nM) (Figure MP*15) are quite similar,[368] the phenylpropyl P_1' group possesses an excellent bond vector at the 4-position to access the remainder of the large S_1' subsite. Thus, the inhibitor **MP-16**, in which the biphenylpropyl group nearly fills the S_1' "tunnel" as shown in Figure MP*16, has a K_i against MMP-3 at pH 7.5 of 2 nM and is over 100-fold more potent than **MP-15**, which bears the parent phenylpropyl P_1' group.[368] In the **MP-16**:MMP-3 crystal structure, the propyl linker is fully extended, and the dihedral angle between the phenyl rings is 30°, which is near the mean value of 26° (range of 0.5° to 58°) in 53 examples of relevant biphenyl systems found in the Cambridge Structural Database.[381] Favorable interactions of the biphenylpropyl group with the protein include the following: (1) the edge of the first phenyl ring is in close contact (3.1 Å) with the carbonyl oxygen of Tyr-237; (2) the edge of the second (terminal) phenyl ring contacts (3.4 Å) the carbonyl oxygen of Ala-234; (3) extensive favorable VDW contacts with the side chains of Leu-215, Val-216, His-218, Leu-235, Tyr-240, Leu-243, and Phe-249; (4) close contact between the terminal phenyl ring and the imidazole ring of His-241. The latter interaction is particularly interesting, as the imidazole ring is positioned over the face of the phenyl ring, such that Cϵ and Nϵ of His-241 are 3.1 Å from C-3′ and C-4′ of the phenyl ring, with a ∼40° angle between the planes of the two rings. Since His-241 has been observed to readily adopt other orientations (as part of the flexible loop),[368] the observed interaction most likely is attractive. A reasonable explanation is that the imidazole is protonated, and a favorable cation-π electrostatic interaction is occurring.[382]

Although **MP-17** (MMP-3 K_i = 14 nM) is 7-fold less potent than **MP-16**, most if not all of the 1.2 kcal/mol difference in the free energy of binding can probably be attributed to the substantially greater entropic penalty incurred by torsional restriction of the flexible butyl group compared to the rigid phenyl ring. This flexibility is illustrated by alternative conformers of the butyl chain in the two independent molecules in the asymmetric unit of the **MP-17**:MMP-3 complex: in molecule 1, a gauche torsion is observed for the penultimate C−C bond, whereas the fully extended conformer is present in molecule 2.[368] The fluorenyl group in **MP-18** further rigidifies the biphenyl system but binds with essentially equal affinity as **MP-16** to MMP-3.[368] In this case, the entropic advantage of inhibitor rigidification and the additional space-filling provided by the C_9 methylene group of the fluorenyl ring system may be compensated for the strain induced by the observed main chain movement of the His-214 residue (the side chain is disordered) to accommodate the planar ring system.

One of the common assumptions in structure-based design of enzyme inhibitors is that a structural unit that binds tightly, such as the P_1' biphenyl group in the present example, will exhibit a single mode of binding throughout a series of inhibitors. It is perhaps surprising, therefore, that two conformations of the propyl linker have been observed in the cocrystal structures of biphenylpropyl-containing inhibitors with MMP-3: **MP-16** and **MP-18** have the extended conformation, whereas **MP-12** and **MP-16** possess a gauche conformation about the central bond of the propyl linker. In addition, the cocrystal structure of **MP-16** with MMP-7 shows the gauche conformer of the propyl linker, while the same inhibitor binds, at least in the crystal, to MMP-3 with an extended propyl linker. Furthermore, the extended conformation of the propyl linker of **MP-15** is observed in its complex with MMP-7.

The structures of the **MP-15**:MMP-7 and **MP-16**:MMP-7 complexes illustrate another level of conformational flexibility of the S_1' loop. As mentioned above, the side chain of Tyr-214 blocks and forms the floor of the S_1' pocket in the cocrystal structures of MMP-7 with inhibitors possessing a small isobutyl group as the P_1' sidechain. From an overlay of the **MP-16**:MMP-3 and **MP-4**:MMP-7 structures, it is clear that Tyr-214 should absolutely block binding

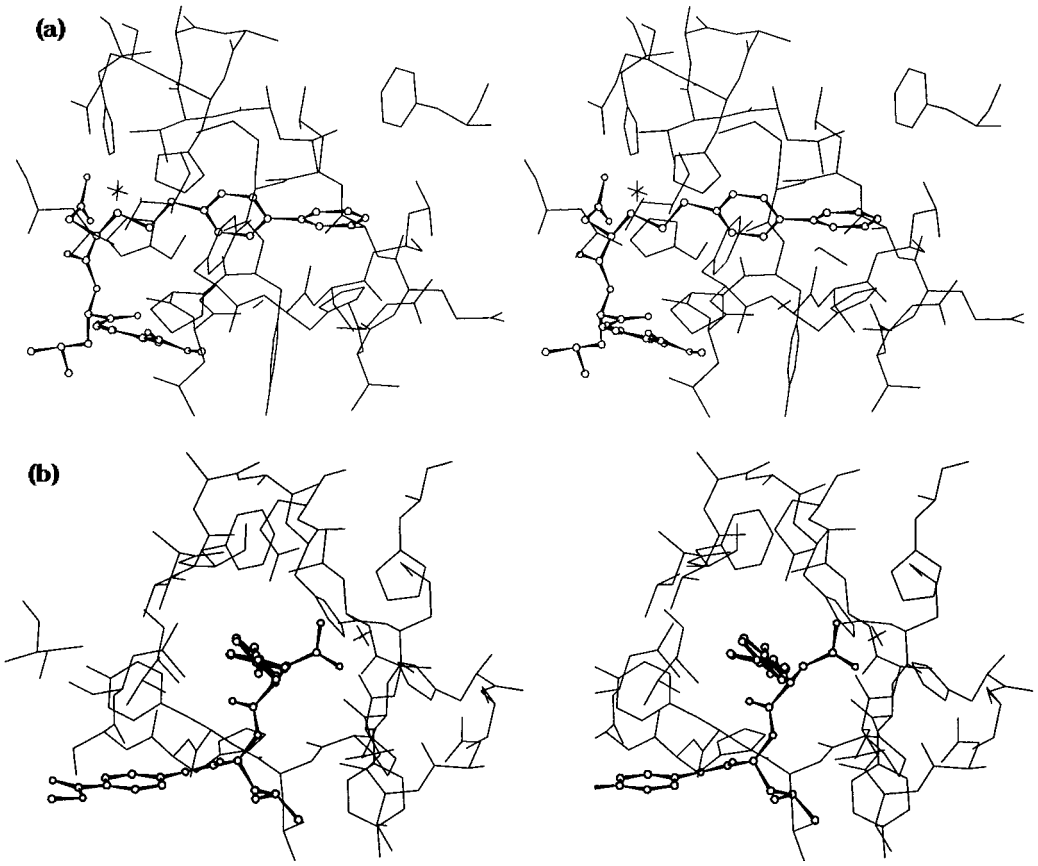


Figure MP*16. Orthogonal stereoviews of **MP-16** (ball and stick) bound to MMP-3.

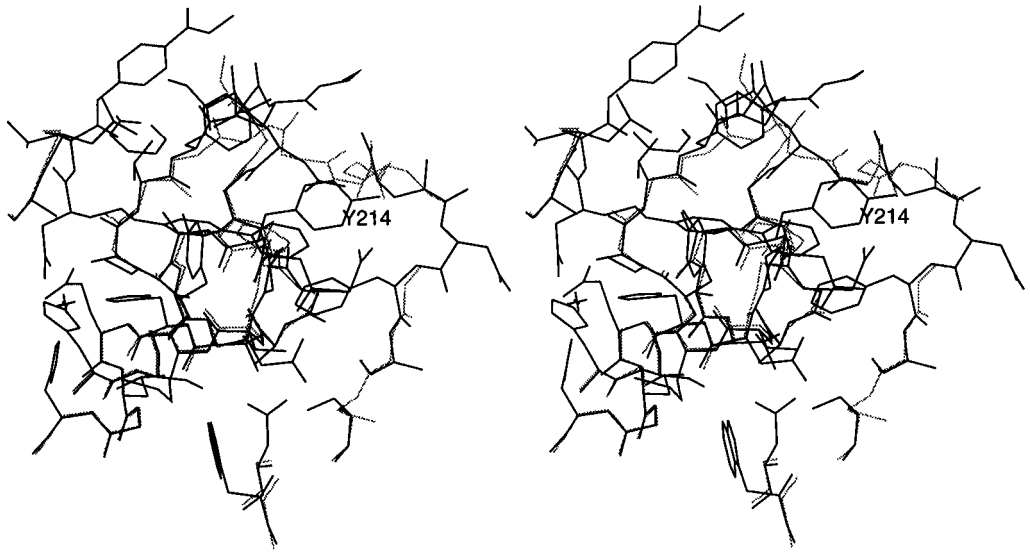
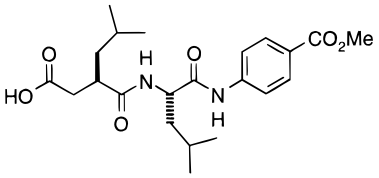


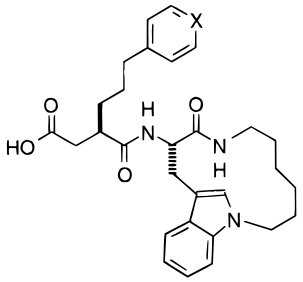
Figure MP*17. Stereoview of the **MP-16**:MMP-7 complex (black) superimposed on the protein structure (backbone atoms plus Tyr-214 side chain) of the **MP-4**:MMP-7 complex (gray), illustrating the movement of the Tyr-214 side chain and the S_1' loop.

of the very large biphenylpropyl group of **MP-16**, leading to the expectation that binding of **MP-16** to MMP-7 should be many kilocalories per mole weaker than the binding of the corresponding analog **MP-19** with isobutyl as the P_1' group. In fact, the binding of **MP-16** is only about 10-fold weaker than its P_1'-isobutyl analog **MP-19**.[347] The cocrystal structure of **MP-16**:MMP-7 shows that a dramatic conformational shift in the S_1' loop (residues 224−233) has created a large S_1' "tunnel" for the biphenylpropyl group that closely resembles the S_1' subsite seen normally for MMP-3 and MMP-8, with the side chain of Tyr-214 now tucked along the side of the subsite (Figure MP*17). Interestingly, the crystal structure of the **MP-15**:MMP-7 complex also shows the same "open" S_1' conformation.

The available data permits an approximate deconvolution of the binding energetics of small and large P_1' groups to MMP-7. If we consider that, in the region where the isobutyl and phenylpropyl groups are interacting, the S_1' subsite of MMP-7 in its "open" loop conformation very closely resembles the S_1' subsite of MMP-3, then the 1.0 kcal/mol increase in binding affinity of **MP-15** over **MP-19** for MMP-3

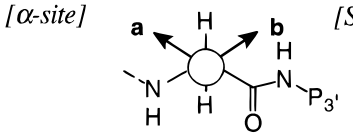


MP-19



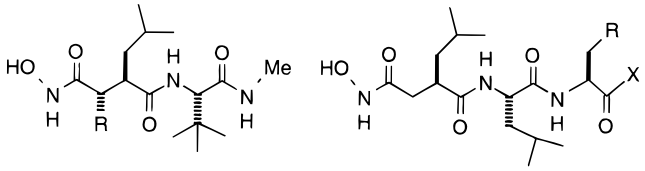
MP-20: X=N
MP-21: X=CH

Figure MP*18.



a (MP#'s): 1(MMP-8), 14, 15, 16, 17, 28, 24, 26
b (MP#'s): 1(MMP-1), 6, 10, 11

Figure MP*19. Torsional orientations of the P_2' side chain in class I inhibitors.



MP-22: R = H
MP-23: R = OH
MP-24: R = H, X = OEt
MP-25: R = Ph, X = NH_2

Figure MP*20.

should also apply to MMP-7 in the hypothetical case where both inhibitors bind to the open conformation. Experimentally, however, **MP-19** binds more tightly to MMP-7 than **MP-15** by 2.5 kcal/mol. Thus, in the complex of **MP-19** (or other inhibitors which bear isobutyl at P_1') and MMP-7, the closed conformation can be estimated to be ~3.5 kcal/mol more stable than the open loop conformation. It is worth noting that **MP-16** is only 8-fold more potent than **MP-15** against MMP-7, even though **MP-15** has already "paid" the energetic cost of inducing the open loop conformation. This relatively modest factor of 8 is much less than that observed for MMP-3 (*vide supra*), but is similar to the potency ratio observed for this pair of inhibitors for other MMPs (e.g., MMP-2, MMP-9, MMP-13) which possess a large S_1' subsite. These data suggest that the S_1' subsite of MMP-3 interacts with the terminal phenyl ring of **MP-16** in a special way not available to the other MMPs. The most likely candidate for this special interaction is the aforementioned contact with imidazole ring in the side chain of His-241, since all other MMPs possess other amino acid residues at this site.

We can conclude our discussion of the characteristics of binding to the S_1' subsite by returning to MMP-1, in which Arg-214 plays a similar role as Tyr-214 in MMP-7. The question naturally arises: can the S_1'-loop of MMP-1 also rearrange to provide an open conformation? The crystal structure of **MP-20** (Figure MP*18) with MMP-1 (K_i = 1900 nM) shows that the side chain of Arg-214 can be displaced to form a larger S_1' subsite, with only modest movements of the loop. The phenylpropyl analog **MP-21** presumably binds to MMP-1 in a very similar manner as **MP-20**, allowing once again the estimation of conformational energetics. In this case, the differential in binding free energy to MMP-1 is 1.4 kcal/mol, favoring **MP-4** over **MP-21**. Essentially the same result is obtained with the inhibitor pair **MP-15** and **MP-19**, although these are much less potent inhibitors against MMP-1 due to effects at P_2'/P_3' (*vide infra*). By using the same 1.5 kcal/mol value as above for the intrinsic preference of the open form for phenylpropyl over isobutyl, a value of about 3 kcal/mol is derived for the preference for the closed loop conformation in the **MP-4**:MMP-1 complex. In the case of MMP-1, the biphenylpropyl analog of **MP-21** binds with lower affinity than **MP-21** itself. An overlay of the **MP-20**:MMP-1 and **MP-16**:MMP-3 structures suggests that the side chain of Arg-241 may not be able to escape unfavorable close contacts with the terminal phenyl ring of the biphenyl system as readily as the side chain of Tyr-214 in MMP-7.

e. Binding Interactions at the S_2' and S_3' Subsites. The S_2' subsite in the MMPs is a solvent-exposed cleft, whose surface is largely defined by the side chains of residues 179 and 180 along the "top" face and the side chain of residue 239 on the "bottom" face. Although in certain MMPs, some of the side chains are polar (see Table MP*2), it has generally been found that hydrophobic side chains at P_2' provide the highest binding affinities. Within a wide range of hydrophobic groups, however, only small differences in binding are observed. For example, leucine, cyclohexylalanine, valine, cyclohexylglycine, phenylalanine, tyrosine, *O*-methyltyrosine, tryptophan, and *tert*-butylglycine can all be incorporated at the P_2' position of class I inhibitors with generally no more than a 5-fold span of affinity for most if not all of the MMPs. In a series of cocrystal structures, β-unbranched side chains (i.e., from leucine, phenylalanine, tyrosine, and tryptophan) have been observed in two torsional orientations about the $C_\alpha-C_\beta$ bond, directing the γ-carbon either toward the α-subsite or toward the S_3' subsite (Figure MP*19).

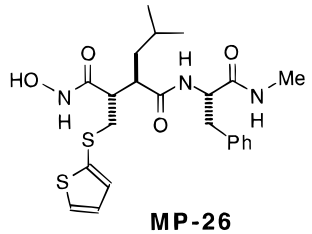
Perhaps the most interesting of the simple, hydrophobic P_2' residues, from both the medicinal and molecular recognition perspectives, is *tert*-butylglycine, which confers nearly equal affinity as leucine, phenylalanine, and tryptophan. Importantly, incorporation of *tert*-butylglycine appears to consistently enhance bioavailability, as illustrated by two orally available class I hydroxamates **MP-22** (Ro31-9790) and **MP-23** (BB-2516, marimastat)[355] (Figure MP*20). Although no cocrystal structures of *tert*-butylglycine-containing dipeptide inhibitors are available, molecular modeling together with the crystal structure of the **MP-12**:MMP-3 complex suggest that the *tert*-

butyl group is probably inferior with respect to VDW interactions compared to other, more extended side chains, and may even induce some strain due to close contacts with the carbonyl oxygen of Pro-238. Factors which effectively compensate for these penalties most likely include more facile desolvation of the adjacent peptide linkages and a conformational effect that constrains the nearby flexible torsions to that found in the bound state (i.e., preorganization).

The active site of the MMPs essentially ends at the S_3' subsite, and the most important feature of this relatively ill-defined region is a hydrophobic wall formed by the Leu-181 side chain and Tyr-240 that separates the S_1' and S_3' subsites. For the dipeptide class I inhibitors, SAR at P_3' basically divides into two categories: simple alkyl groups, most usually methyl, and aryl and heteroaryl groups, often substituted at the 4-position. Early work established that replacement of the P_3' residue of a tripeptide inhibitor with a methyl group resulted in equal if not greater activity.[383] This SAR is consistent with the crystal structures of **MP-24**:MMP-1 (PDB entry: 1TCL)[384] and **MP-25**:MMP-3,[385] which show that the peptide chain turns toward solvent at the P_3' position. It is worth noting that the reported NMR solution structure of the **MP-25**:MMP-3 complex (PDB entry: 1UMS and 1UMT)[386] differs substantially, especially in the active-site region, from all crystallographic structures and illustrates the potential weaknesses of the NMR method for determining the structure of protein−ligand structures.

For the collagenases and gelatinases, little or no gain in binding affinity can be gained by P_3'-substitution larger than a methyl group, with the exception of certain macrocyclic derivatives to be discussed below. Small aryl groups are tolerated by the collagenases at P_3' but extension from the 4-position results in a modest loss of activity, presumably due to close contacts with the aromatic ring of Tyr-210, which is conserved and forms the "end" of the S_3' subsite in the collagenases and gelatinases. For example, the 4-pyridyl group at P_3' is about equipotent with methyl against MMP-1, but the 4-carbomethoxyphenyl group is about 2-fold less potent.[368] In contrast, aryl substitution at P_3' leads to a large increase in binding affinity to MMP-3 and MMP-7, where residue 210 is threonine and isoleucine, respectively. For example, the aforementioned 4-carbomethoxyphenyl group at P_3' can result in as much as 40-fold increase in binding to MMP-3 and MMP-7 relative to a P_3' methyl group.[368] Examination of the cocrystal structures of **MP-16** with MMP-3 and MMP-7 reveals that the P_3' aryl ring is packed against the "wall" and that a reasonable, although not optimal, edge-to-face interaction is occurring with the ring of Tyr-240.

Before structural information became available for the MMPs, researchers at Roche found that cyclization of the P_2' and P_3' groups was well-tolerated, provided that the ring size was large enough to permit the *trans* amide to be the predominant rotamer.[383,387] Workers at Syntex extended these findings and reported that macrolactams derived from tryptophan at P_2', as in **MP-2** through **MP-6**, increase activity about 10-fold relative to that of acyclic analogs.[360] It is likely that preorganization is largely responsible for this effect, since the cocrystal structures of these inhibitors (e.g., **MP-2, MP-3, MP-4, MP-5, MP-7, MP-20, MP-21**) reveal that the indole ring spatially coincides very closely with that of the acyclic inhibitor **MP-7**.



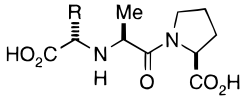
MP-26

Figure MP*21.

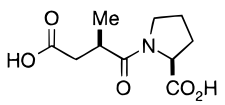
f. Binding Interactions at the α Subsite. In the case of class I hydroxamate inhibitors, stereospecific substitution at the α-position with either substituted methylene (i.e., R-CH_2) or hydroxy appears to enhance potency against most if not all MMPs. A good example of the former case is **MP-26** (batimastat, BB-94, Figure MP*21), which possesses subnanomolar K_i values for all tested MMPs.[388] Analysis of the crystal structure of **MP-26** bound to MMP-8 (PDB entry: 1MMB)[389] identifies several potential factors that could lead to enhanced binding: (1) preorganization of the $\alpha-\beta$ dihedral to the bound state, since the α-CH_2 is anti (178°) across this torsion with the CH_2 of the P_1' isobutyl group in the bound conformation; (2) preorganization of the inhibitor through hydrophobic collapse,[125] since there is a close VDW contact between the CH_2S linker and C-3 of the thiophene with the phenyl ring at P_2' (e.g., 3.3 Å from C-3 of thiophene to C-4 of phenyl); (3) an edge-face interaction of the imidazole ring of His-228 with the thiophene ring. In addition, substitution of a nonpolar group at the α-position would be expected to decrease the energetic costs of desolvating the polar hydroxamic acid function. Despite these interactions, substitution of the thienylthiomethyl group of **MP-26** with a simple methyl group results in ≤5-fold loss in binding affinity to the MMPs tested. The cocrystal structure of **MP-26** bound to MMP-1 determined at Agouron is essentially the same,[354] with slight deviations in the position of the α- and P_2'-side chains. From these structures, it is clear that the α-substituent in class I inhibitors can indeed access the same general space as a P_1 side chain of a peptide substrate, although the approach is from a nearly perpendicular direction. Also noteworthy is the binding mode of **MP-26** to atrolysin C (EC 3.4.24.42), a snake venom metalloproteinase with substantial structural similarity to the MMPs, which has been determined crystallographically (PDB entry: 1DTH) to be fundamentally different, with the thienylthiomethyl side chain occupying the S_1' pocket.[390]

3. Class II Pseudopeptide Inhibitors

a. Overview of SAR. Inhibition of MMPs with class II carboxylates have been extensively explored primarily due to the great success that members of this structural class have enjoyed in inhibiting ACE. In studies leading to enalaprilat (**MP-29**, Figure MP*22), a key observation was that introduction of



MP-27: R=H
MP-28: R=Me
MP-29: R=$(CH_2)_2Ph$



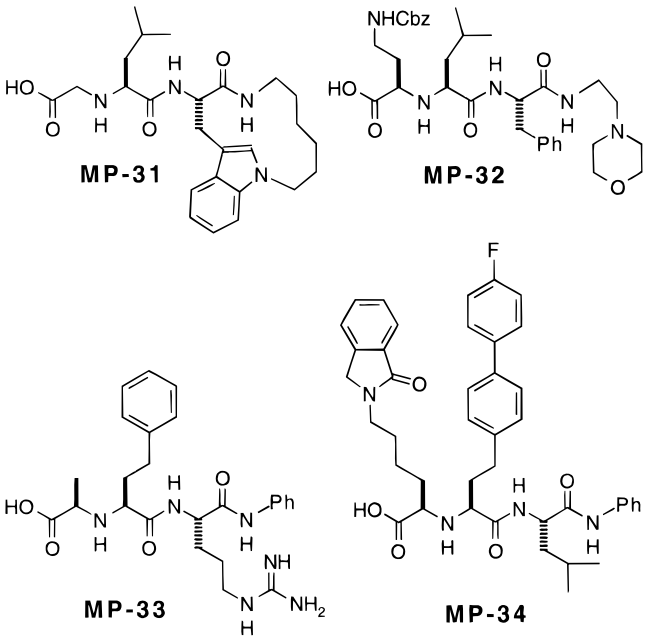
MP-30

Figure MP*22.

a suitable alkyl group at the α-position to the carboxylate improved potency dramatically: the α-unsubstituted version **MP-27** was of comparable potency to the analogous class I carboxylate **MP-30**, but introduction of the α-methyl group in the *S* configuration (e.g., **MP-28**) improved potency against ACE 50-fold.[338] Similarly, α-unsubstituted *N*-carboxymethyl dipeptides inhibit MMPs with comparable potency to the corresponding class I carboxylates, but in the case of the MMPs, α-substitution increases potency only modestly (2−10-fold for a methyl group),[391] and interestingly, the opposite *R* configuration at the α-position is preferred. Simple derivatives containing standard substrate-like side chains inhibit only in the micromolar concentration range.[391] Class II glutaramide inhibitors (Z = CH_2) appear to be of similar potency relative to the *N*-carboxyalkyl inhibitors (Z = NH),[391] which is also true for α-unsubstituted ACE inhibitors. Furthermore, *N*-phosphonoalkyl dipeptide inhibitors (Z = NH, ZBF = PO_3H_2) appear to have similar or slightly enhanced potency relative to the corresponding carboxylate inhibitors.[392] Introduction of certain large α-substituents has led to very potent class II carboxylate inhibitors,[391,393,394] but, in general, these relatively high molecular weight inhibitors have lacked suitable bioavailability. A very recent report that glutaramide inhibitors bearing 4′-substituted biphenylethyl groups at P_1' are potent, selective, and orally bioavailable MMP inhibitors is especially intriguing because the differences in binding of *N*-carboxyalkyl (Z = NH) and glutaramide (Z = CH_2) inhibitors are analyzed in some detail (*vide infra*).[395]

b. Interactions of Class II inhibitors at the Zinc Site. Coordinates for three cocrystal structures are available of MMPs complexed with class II inhibitors, which are all *N*-carboxylalkyl dipeptides (i.e., Z = NH, ZBF = CO_2H): **MP-31**:MMP-1,[396] **MP-32**:MMP1 (PDB entry: 1CGL),[349] and **MP-33**:MMP-3 (PDB entry: 1SLN)[375] (Figure MP*23). These structures from three different research groups exhibit nearly identical binding geometries of the inhibitor, and clearly illustrate that both the carboxylate and the amino group are involved in key interactions at the catalytic site. In all three cases, the carboxylate binds to the zinc in a monodentate manner, with a Zn−O distance of 1.75−2.11 Å, resulting in a near-tetrahedral coordination geometry. The other carboxylate oxygen is 2.6−3.0 Å from the zinc and is solvent exposed. In contrast to class I inhibitors, Glu-219 does not H bond to the zinc-bound oxygen of the inhibitor (O−O distance of 3.1−3.7 Å), but rather to the nitrogen of the inhibitor. This nitrogen is also within H-bonding distance of the carbonyl oxygen of Ala-182.

While the protonation state in the active site of the class I carboxylate−MMP complexes appears unambiguous, the protonation state in the case of class II *N*-carboxyalkyl inhibitors is less clear. Three reasonable situations can be proposed, which are illustrated in Figure MP*24. In the first, the amine group of the inhibitor is protonated and Glu-219 is unprotonated: a strong H bond to Ala-182 C═O and a salt bridge to Glu-219 Oϵ1 are favorable consequences, but the introduction of another positive charge at close proximity (4.5−4.7 Å) to the zinc would appear to be unfavorable. The second situation has the same H bonds as in first, but in this case, the amine is unprotonated (neutral) and Glu-219 Oϵ1 is protonated. The difference between these two cases is only the precise position of the proton between Glu-219 and the nitrogen of the inhibitor. An NMR determination of the structure of **MP-26** bound to MMP-3 has been reported which could have provided experimental evidence concerning the protonation state of the nitrogen of the inhibitor, but there was no mention of data to resolve this point.[397] The third case is analogous to the first case, except that Glu-219 is protonated on Oϵ2. In view of the loss of the favorable salt bridge and the buildup of positive charge, it seems likely that this case would only be relevant at lower pH.



MP-31

MP-32

MP-33

MP-34

Figure MP*23.

Although coordinates are not yet available (PDB entry: 1HFS, on hold until February, 1998), the Merck group has reported the structure of the glutaramide inhibitor **MP-34** bound to MMP-3,[395] which serves to highlight the sensitivity of binding geometry to the the substitution of interacting polar functionality with nonpolar functionality, in this case, substituting Z = CH_2 for Z = NH. Since the CH_2 group in MP-34 can no longer H bond to either the Ala-182 C═O or Oϵ1 of Glu-219, it is essentially forced to move to (at least) a VDW contact distance: reported distances are 3.59 Å to the carbonyl oxygen of Ala-182 and 5.22 Å to Oϵ1 of Glu-219.[395] This movement is accompanied by a change in the carboxylate−zinc ligation geometry, which is now much more like that of class I carboxylate inhibitors: unsymmetrical, bidentate coordination (1.8 and 2.6 Å) with a strong

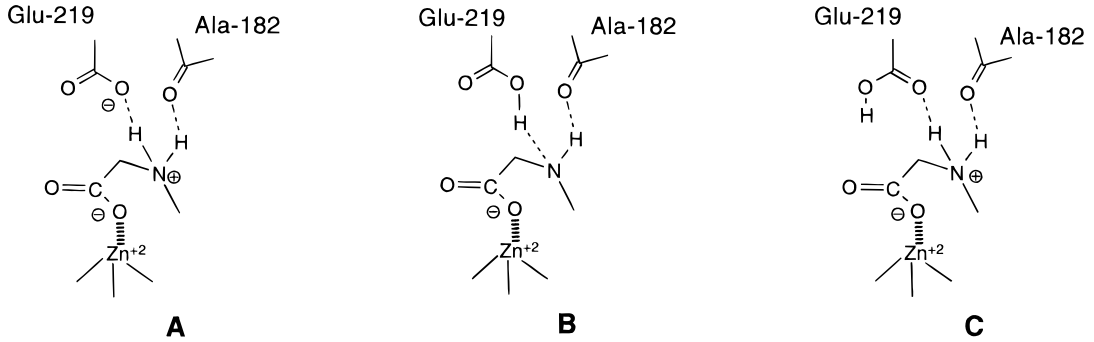


Figure MP*24. Possible protonation states for the complex of MMP with carboxyalkylamine inhibitors.

2.7 Å H bond of the "inner" carboxylate oxygen to Oϵ1 of Glu-219, which is presumed to be protonated.[398]

c. Backbone Interactions of Class II Inhibitors with MMPs. In general, the peptide backbone of class II dipeptide inhibitors fulfills the same role, that of completing a β-strand-like structure within the MMP active site, as observed for the class I inhibitors discussed previously. In all three cocrystal structures of the *N*-carboxyalkyl inhibitors, the class II backbone is shifted ~1.0 Å *toward* the zinc relative to the backbone of class I carboxylates. This repositioning of the backbone may or may not lead to the observed weakening of the P_2' NH to O=C(Pro-238) H bond, which ranges from 3.4 to 3.8 Å in the three structures. The perturbation results from movement of the protein backbone up to ~1 Å, beginning at Tyr-237 and extended into the flexible S_1' loop discussed previously. Despite this movement, the other three backbone H bonds are uniformly well-maintained. On the other hand, the backbone position of the glutaramide **MP-34** in its complex with MMP-3 is reported to be shifted almost 1 Å away from the zinc relative to **MP-33**,[395] and so appears to be more similar in positioning to class I inhibitors. The P_2' H bond to O=C(Pro-238) is this case is 3.36 Å, which is intermediate between the two extreme cases of class I inhibitors and class II *N*-carboxylalkyl inhibitors.

d. Binding Interactions in the S_1' Subsite. Once again, the general orientation of the P_1' side chain in class II *N*-carboxyalkyl inhibitors resembles that of class I inhibitors, but some small differences do occur as a result of the altered vector from the backbone methine carbon to the first CH_2 of the P_1' group. In the case of P_1' isobutyl, as in **MP-31** and **MP-32**, slight torsional adjustments result in the terminal methyl groups being in essentially the same position as class I inhibitors in their respective complexes with MMPs. In the case of a phenethyl group, overlay of the class II inhibitor **MP-33**:MMP-3 structure with the class I inhibitor **MP-14**:MMP-3 structure shows that although the positions of C_4 of the phenyl ring closely correspond, the C_4−H bond vector is significantly different in that the class II bond vector appears better disposed to access the remainder of the large S_1' subsite. SAR from the Merck group has established that simple alkyl substitution at the 4-position does in fact enhance potency;[391] no analogous data has been reported for 4-substitution of P_1'-phenethyl-containing class I inhibitors.

As mentioned above, the *N*-carboxylalkyl inhibitor **MP-33** and the glutaramide inhibitor **MP-34** have significantly different binding geometries to MMP-3, and the details of the interaction of the P1′ substituent with the S1′ subsite also show some interesting differences. In particular, the conformation of the P1′ phenylethyl group in **MP-33** is ttg^- whereas the corresponding torsional arrangement in **MP-34** is g^+tg^-. The combination of backbone shifts and torsional adjustments result in a very similar positioning of the corresponding aryl rings in the two complexes in terms of filled volume, but the bond vectors from the 4-position are not identical, with the consequence that biarylethyl substitution is considerably more favorable in the glutaramide series than in the *N*-carboxylalkyl series.[398] These studies provide yet another example wherein the availability of crystallographic information on several related inhibitors has served to establish unexpected differences in the binding mode which have important implications for the design of improved inhibitors.

e. Binding Interactions at the S_2' and S_3' Subsites. The general trends described for class I inhibitors appear to hold true for class II inhibitors, since the shift in the backbone in class II inhibitors relative to class I inhibitors has little impact on interactions of the P_2' and P_3' side chains with the solvent-exposed S_2' and S_3' subsites. In fact, overlay of the **MP-7**: MMP-7 and **MP-31**:MMP-1 structures shows that the sulfodiimine ZBF of **MP-7** positions the peptidic backbone in virtually the same position as in the *N*-carboxyalkyl inhibitor **MP-31**.

f. Binding Interactions at the α-Subsite. As mentioned above, α-substitution in ACE class II inhibitors resulted in large potency enhancements, while in MMP inhibitors, the opposite configuration is required and improvement in potency is much more modest. The structure of **MP-32** bound to MMP-1 (PDB entry: 1CGL) provides an explanation for these observations (Figure MP*25).[349] First of all, it is clear that inverting the stereochemistry at the α-carbon to that of the ACE inhibitors would lead to severe VDW contacts with the protein, especially the α- and β-carbons of His-183. A consequence of the required stereochemistry, however, is that instead of the extended conformation that would result from the "ACE configuration", the inhibitor adopts a folded conformation in which a distorted double-gauche-pentane-like arrangement is present about the $C_\beta H_2$−C_αH−NH−CH−C(=O) unit. As a result, C_β and the carbonyl carbon of the P_1' leucine residue

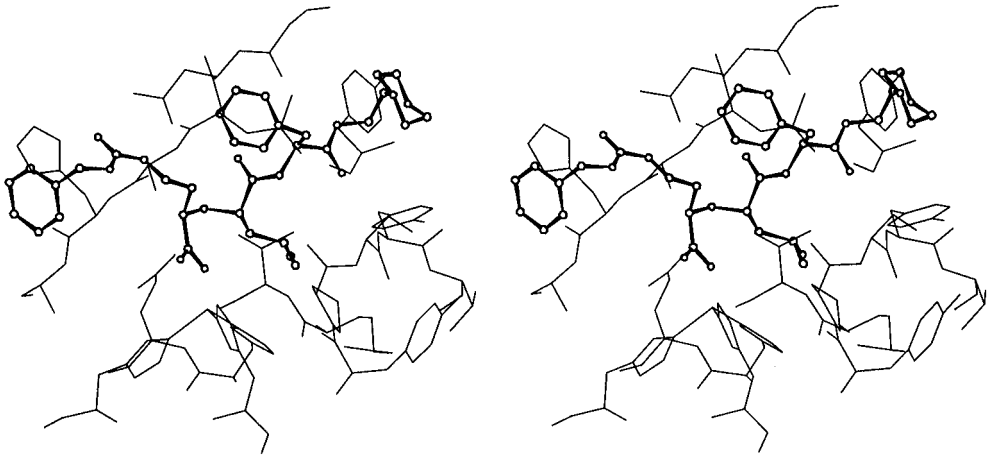
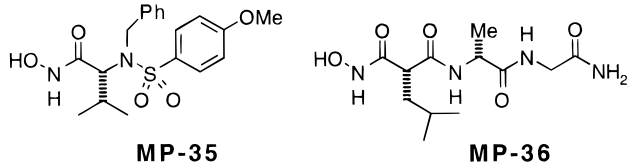


Figure MP*25. Stereoview of **MP-32** (ball and stick) bound to MMP-1.



MP-35 **MP-36**

Figure MP*26.

have a close 3.3 Å contact, and the $C_\alpha H—NH—CH—C(=O)$ dihedral is −108°, which suggests that the inhibitor binds in a significantly strained conformation. Furthermore, the carbamate group in the α-substituent is bound in its higher energy *cis* rotomer. Compensating for the conformational penalty are a number a favorable VDW and H-bonding interactions of the α-substituent with the protein: (1) the carbamate group H bonds to both the Asn-180 side chain and to the imidazole of His-183 that ligates to the structural zinc; (2) the phenyl ring of the Cbz group makes a favorable stacking interaction with the imidazole ring of His-183, occupying approximately the same position as the P_3 side chain of a peptide substrate.

4. Class III Inhibitors

a. Overview of SAR. Inhibition of MMPs by *N*-(arylsulfonyl)-α-amino *N*-hydroxy amides (e.g., **MP-35**) was independently discovered by groups at Ciba-Geigy[399] and at British Biotech[400,401] (Figure MP*26). Only a single malonadiamide derivative **MP-36** has been described in the journal literature,[374] and it binds weakly to MMP-8 (K_i = 33 μM). Nevertheless, the structure of its complex with MMP-8 (*vide infra*) suggests that further optimization of binding, in analogy with the *N*-(arylsulfonyl)-α-amino *N*-hydroxy amides which exhibit a somewhat similar binding mode, should be readily achievable.

b. Interactions of Class III Inhibitors at the Zinc Site. The well-conserved geometry of binding by class I hydroxamate inhibitors could be interpreted to mean that this geometry is intrinsically preferred by hydroxamate ligands, but it may be more likely that the constraints imposed by the peptidic framework of these inhibitors dictates the consensus geometry. Support for the latter view can be obtained by inspection of the structures of the class III inhibitors **MP-35** and **MP-36** bound to MMP-1 and MMP-8, respectively (Figure MP*27). The hydroxamate functions in these two structures are nearly superimposable, despite the substantial differences in the rest of the inhibitor structures. Although these class III hydroxamate groups once again bind in a bidentate manner at zinc, there are subtle differences from the class I structures. Most importantly, the hydroxamate O^- appears to H bond (2.57 Å in the **MP-35**:MMP-1 complex[354]) to the carboxylate Oϵ1 oxygen of Glu-219 rather than to the Oϵ2 oxygen which H bonds in the class I structures. Although it is difficult to assess *a priori* the relative energetics of these two H-bonding arrangements, consideration of the mechanism of peptide hydrolysis, wherein in the Oϵ1 oxygen of Glu-219 presumably transfers a proton to the leaving amine group, suggests the possibility that protonation of the Oϵ1 oxygen may be energetically preferred. The potency of **MP-35** (MMP-3 K_i = 34 nM)[402] is extraordinary in view of its relatively limited interactions with the protein (*vide infra*) when compared to class I peptidic inhibitors of similar potency.

c. Backbone Interactions of Class III Inhibitors. The nonpeptidic class III inhibitors make only one well-defined "backbone" H bond to the MMPs, in which the NH of Leu-181 is the donor and one of the sulfonyl oxygens in **MP-35** and the malonyl amide carbonyl oxygen of **MP-36** serve as the acceptors. The position of the acceptor oxygen is similar to that observed for class II inhibitors, although in the case of **MP-36**, the C=O and N—H bond vectors are less colinear than in the case of class I or class II inhibitors. In the structure of **MP-35**:MMP-1,[354] the carbonyl oxygen of Pro-238, which serves as a "backbone" H bond acceptor in class I and class II inhibitors, undergoes a small but significant displacement relative to its position in class I and class II inhibitors in order to avoid unacceptable nonbonded contacts with the CH_2 group of the *N*-benzyl substituent. This provides additional evidence for the flexibility of this region of the protein. In the case of **MP-36**:MMP-8 complex,[374] the position of the Pro-238 carbonyl is essentially identical to that found in the **MP-26**:MMP-8 structure discussed above, despite a 3.1 Å close contact with the methyl group of the alanine residue of **MP-36**. Also, it should be mentioned that the carbonyl oxygen of the Ala residue of **MP-36** is positioned to interact with the NH of Tyr-240, although the distance (3.4 Å) and the unfavorable

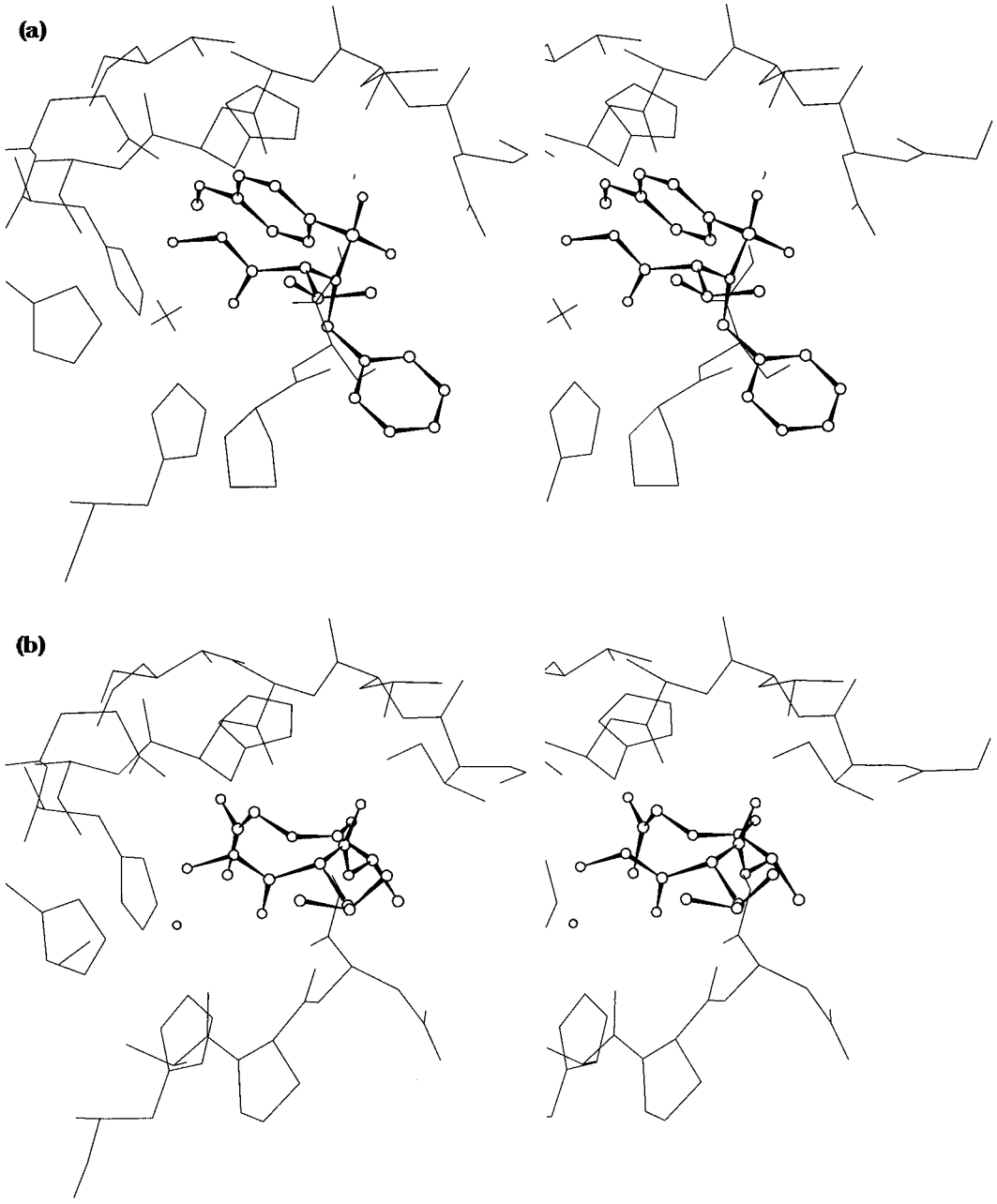


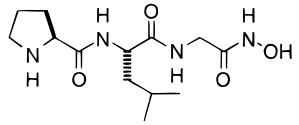
Figure MP*27. Stereoviews of (a) **MP-35** (ball and stick) bound to MMP-1 and (b) **MP-36** (ball and stick) bound to MMP-8.

bond angles indicate that the interaction is probably not significantly stabilizing to the complex. Similarly, a very weak interaction (3.65 Å) between the NH of the Ala residue of **MP-36** and the carbonyl oxygen of Pro-238 is not likely to contribute significantly to binding.

d. Interactions of Class III Inhibitors at the S_1' Subsite. As shown in Figure MP*27a, the 4-methoxyphenyl ring of **MP-35** is binding to the S_1' pocket. In comparison to the interactions of the "optimal" P_1' isobutyl group in class I MMP-1 inhibitors, the aromatic system appears to be an even better fit to the size and shape of the S_1' pocket of MMP-1, making extensive VDW contacts at close to optimal distances. Since the limited SAR available suggests that the 4-methoxy group is specifically advantageous over other substituents, there may also be an electrostatic component involving the relatively large dipole across the aromatic ring that is generated by the methoxy donor and the sulfonyl acceptor.

The inhibitor **MP-36** was envisioned to be a pseudopeptidic inhibitor, in analogy to similar hydroxamate inhibitors of thermolysin, but the cocrystal structure with MMP-8 clearly showed that the intended P_2' and P_3' residues actually were binding in the S_1' subsite. As explicitly acknowledged by the authors,[374] replacement of the "peptidic tail" of **MP-36** with non-peptidic groups designed to interact more favorably with the S_1' subsite should lead to inhibitors with much greater potency than **MP-36**.

e. Binding Interactions at the S_2' and S_3' Subsites. Neither **MP-35** nor **MP-36** extend into the S_2' or S_3' subsites, and no information has been published that suggests that extensions into these binding domains are advantageous.



MP-37

Figure MP*28.

f. Binding Interactions at the α-Subsite. The isopropyl group in **MP-35** and the isobutyl group in **MP-36** clearly occupy the same general position and make similar VDW contacts as the α-groups in class I and class II inhibitors, although the precise bond vectors which attach the α-substituents to the inhibitor backbone are different in the three cases.

5. Inhibition of MMPs by LHS Inhibitors

The complex of **MP-37** with two different truncated forms of MMP-8 (pdb:1JAN and 1JAP; Figure MP*28) has been reported, and shows essentially identical binding modes.[374] The hydroxamate group of LHS inhibitor **MP-37** also binds in a bidentate manner, with a geometry that is essentially a ~90° rotation about the N−H bond of the hydroxamate relative to the class I hydroxamate geometry. With this geometry, H bonding to Glu-219 and Ala-182 is preserved, but the alternative positioning results in a close contact (2.8 Å) between the hydroxamate carbonyl oxygen and the C-2 carbon of the ligating imidazole of His-228. In addition, the coordination geometry around zinc is close to square pyramidal, with the zinc atom slightly (0.7 Å) above the plane of the equatorial ligands and His-222 occupying the apical position.[374] It seems reasonable to assume that the poor affinity (K_i = 19 μM) of **MP-37** for MMP-8, relative to class I hydroxamates is in part due to suboptimal hydroxamate binding.

The binding of the peptide backbone of **MP-37** likely reflects the "unprimed" binding interactions of peptide substrates and involves the following interactions: (1) the carbonyl oxygen and the NH of the P_2 Leu residue H bond (2.9 Å each) to the NH and C═O of Ala-184, respectively; (2) the P_3 proline ring occupies a small hydrophobic pocket formed by the aromatic side chains of His-183 and Phe-185; (3) the P2 Leu side chain makes weak VDW contacts with the protein surface. The P_1 NH and side chain (absent in **MP-37**) are both solvent exposed, as is the N-terminal nitrogen of the P_3 Pro residue.

D. Summary and Perspective

Taken together with the extensive SAR that has been reported, structural studies on the complexes of MMPs with small molecule inhibitors has served to clearly distinguish the binding elements that are primary in importance from those that are secondary: (1) The nature of the zinc-binding function is of outmost importance, and the structural studies shed light on why neutral ZBFs, most particularly the hydroxamic acid function, are advantageous over ionized ZBFs such as carboxylates and phosphinates. (2) The primary specificity interaction site has been clearly identified as the S1′ subsite, and intriguing examples of protein conformational flexibility have been delineated in the response of the S1′ site to changes in the size of the P1′ group of inhibitors. (3) Substrate-like RHS inhibitors bind with the P1′/P2′ and P2′/P3′ amide groups H bonding to the protein in a β-sheet-like fashion. Of the four H bonds normally observed, however, only the one from the NH of Leu-181 to an acceptor on the inhibitor is absolutely conserved; it is worth noting that both structures of **MP-37**:MMP-8, which lacks "prime-side" residues, also show an ordered water accepting a H bond from the NH of Leu-181, while the other three H-bonding donor−acceptors do not have conserved waters within H-bonding distance. One possible reason for these observations is that the NH of Leu-181 is "electrostatically hot" due to complexation of carbonyl oxygen to the nearby structural calcium ion. (4) Binding interactions at the α, S_2', and S_3' subsites play a relatively minor role in determining binding affinity, consistent with their location at the protein−solvent interface.

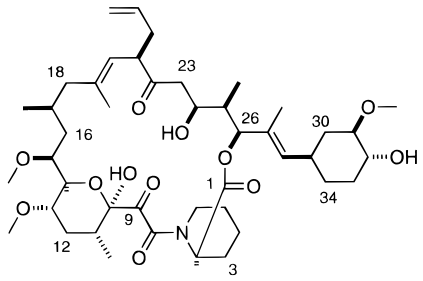
VII. Immunophilins, Immunosuppressive Agents, and Chemical Inducers of Dimerization

A. Introduction

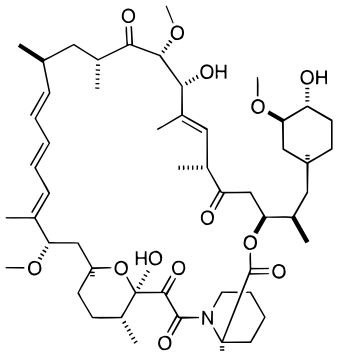
Cyclosporin A (**CSA**),[403] FK506 (a.k.a tacrolimus),[404] and rapamycin (**IM-1**)[405,406] are natural products which have activity as immunosuppressive agents. (See Figure IM*1.) In T-cell culture, cyclosporin A and FK506 inhibit the production of the cytokine IL-2. Rapamycin does not inhibit production of IL-2, but rather it inhibits the cell's response to IL-2. Cyclosporin A is a high-affinity ligand for the peptidyl−prolyl cis−trans isomerase (PPIase) cyclophilin.[407−409] Both FK506 and rapamycin are high-affinity ligands for an unrelated PPIase FKBP12.[410,411] Proteins related to cyclophilin and FKBP12 are referred to as immunophilins.[412,413] Interpretation of a variety of data suggested that the immunosuppressive activities of cyclosporin A,[414] FK506, and rapamycin was not due to their ability to inhibit the PPIase activity of immunophilins (i.e., "loss of function model"), but rather that these natural products in complex with their respective immunophilins acquired a "gain of function".[415]

The "gain of function" hypothesis was confirmed when it was discovered that FK506 in its complex with FKBP12 bound to and inhibited the serine/threonine phosphatase calcineurin.[416] Neither FK506 alone nor FKBP12 alone inhibited calcineurin. In addition, the FK506 complex with various mutants of FKBP12 had reduced affinity for calcineurin.[417−419] It was also found that the cyclosporin A/cyclophilin complex also inhibited calcineurin. Neither cyclosporin A alone nor cyclophilin alone inhibited calcineurin. Thus, the ability of both cyclosporin and FK506 to prevent formation of IL-2 could be attributed to their ability, in complex with a different PPIase enzyme, to inhibit calcineurin.[416] The rapamycin/FKBP12 complex was found to bind to a new protein target, FRAP (FBBP12 rapamycin associated protein), neither rapamycin alone nor FKBP12 alone bound to FRAP.[420]

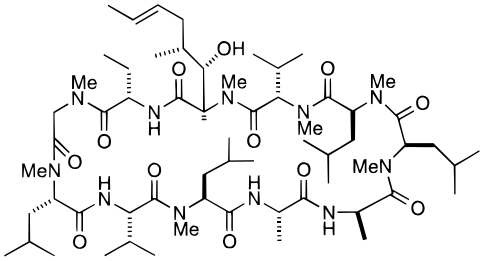
These three natural products were the first characterized examples of ligands which could simultaneously bind to two different proteins. Binding to



FK506



Rapamycin (**IM-1**)



Cyclosporin A (**CsA**)

Figure IM*1. Structures of immunosuppressive natural products.

two different proteins provides a mechanism for bringing together two protein partners that normally have little, or no affinity for each other. These types of compounds have been named "molecular glue inhibitors" [421] and "chemical inducers of dimerization".[422]

In this section we will review the nature of the molecular recognition between FKBP12 and many of its ligands. This will include inhibition of PPIase activity and structural studies on many FK506 analogs which do not inhibit calcineurin. Some of the SAR data will be discussed in terms of the structures of complexes of related ligands and hydrophobic effects. The ternary complex of FKBP12, FK506, and calcineurin and the ternary complex of FKBP12, rapamycin, and the FRB domain of FRAP will be analyzed. These structures provide the first examples of how a small molecule ligand can induce the heterodimerization of two proteins. The final topic will briefly review the competitive inhibition of serine/threonine phosphatases by natural ligands.

B. FKBP12-FK506 Complex and FKBP12-Rapamycin Complex

The structure of the FK506-FKBP12 complex has been determined by x-ray crystallography (PDB entry: 1FKF).[423,424] The pipecolinic binding site of FKBP12 is a large, shallow hydrophobic cleft. (See Figure IM*2.) The binding site occupied by the pipecolinyl ring of FK506 is formed by the convergence of several aromatic groups. These groups are the side chain of Trp-55, which forms the "base" of the binding site, and the side chains of Phe-46, Tyr-26, and Phe-99 which form the "walls". Other hydrophobic side chains of Val-55 and Ile-56 also form the walls of this binding site. Approximately one-half of the FK506 surface is buried at the protein−ligand interface, the other half is solvent exposed in the FKBP12-FK506 complex. Thus, the complex exposes hydrophobic surface of FK506 into solvent. In addition to burying ligand hydrophobic surface there are four ligand−protein hydrogen bonds. The ester carbonyl of the macrocycle forms a hydrogen bond with the backbone NH of Ile-56 (2.8 Å), the amide carbonyl oxygen hydrogen bonds to the side-chain phenol of Tyr-82 (2.8 Å), the C24 hydroxyl forms a solvent-exposed hydrogen bond with the backbone carbonyl of Glu-54 (2.7 Å), and the hemiketal hydroxyl makes a solvent-exposed hydrogen bond to one of the anti-lone pairs of the side chain of Asp-37 (2.8 Å). The C8 amide carbonyl, Tyr-82 hydrogen bond is nearly orthogonal to the carbonyl plane. The pipecolinyl ring is buried in the complex and makes VDW contacts with the four aromatic residues Trp-59, Phe-99, Tyr-26, and Phe-46. The C11 methyl group is buried in a hydrophobic pocket defined by Ile-90, Ile-91, and His-87. This region of FKBP12, from residues Ala-84 to His-94, is quite flexible and can accommodate a variety of different groups. The cyclohexyl group of FK506 lies on a shallow hydrophobic groove of FKBP12 and is partially solvent exposed.

The C9 ketone carbonyl of FK506 makes an unusual interaction with FKBP12. Three aromatic residues of FKBP12 converge to form a small pocket. This pocket is lined by three aromatic ϵ-hydrogens, one each from Tyr-26, Phe-36, and Phe-99. The ketone carbonyl oxygen fills this cavity and makes close contacts with one aromatic CH of each residue (3.4−3.6 Å carbon−oxygen distances). This is likely a favorable electrostatic interaction since three aromatic CH dipoles point into a carbonyl oxygen.[425,426]

The pipecolinyl moiety binds as the trans amide rotomer. The ketone carbonyl is approximately 90° out of the plane of the amide carbonyl. The ester group of the macrolactone is axially disposed on the six-membered pipecolinic ring. (See Figure IM*3.) The out-of-conjugation dicarbonyl system and the axial ester group are a local conformational minima. A conjugated ketone carbonyl would result in unfavorable VDW contact as a result of an $A^{1,3}$ strain mechanism. Similarly an equatorial ester group would also suffer unfavorable VDW contacts. Thus, FK506 contains "conformational locks" which disfavor geometries which are not complimentary to the FKBP12 binding site.

In contrast to protease enzymes which catalyze a reaction which occurs extremely slowly in the absence of an enzyme, peptidyl−prolyl cis−trans isomerization occurs readily in the absence of an enzyme. The proposed mechanism for FKBP12 catalyzed peptid-

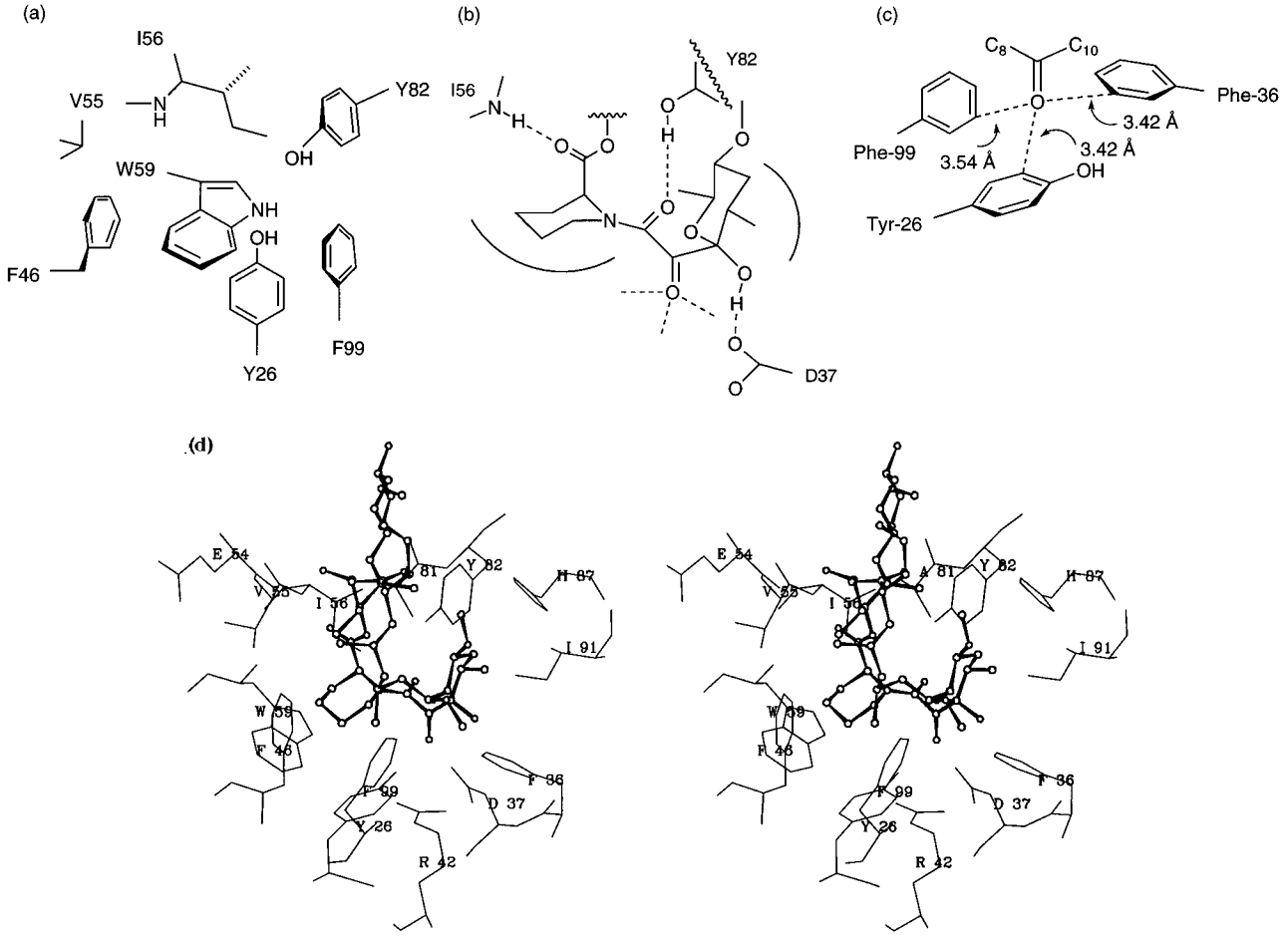


Figure IM*2. (a) Binding site of FKBP12 is a shallow hydrophobic cleft lined by several aromatic residues. (b) Interactions of FK506 "binding domain" with FKBP12. (c) The ketone carbonyl of FK506 binds in a small cavity and interacts with three aromatic C-H bonds. (d) Stereoview of FK506-FKBP12 complex.

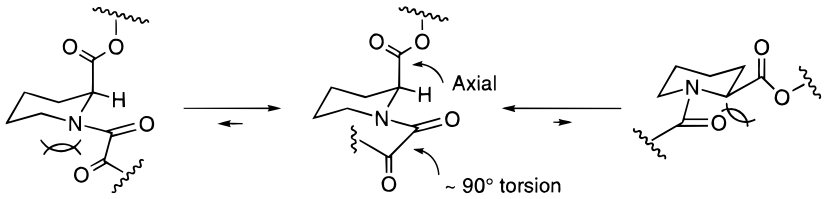


Figure IM*3. The preferred conformation of the pipecolinyl dicarbonyl region of FK506 and related compounds is shown in the middle. The axial substituent on the piperidine ring is preferred since an equatorial substituent would suffer from $A^{1,3}$ strain. The approximate 90° dihedral angle about the dicarbonyl is also preferred since conjugated conformations would suffer from $A^{1,3}$ strain. Thus, the binding conformation is low in energy due to the presence of two local conformational locks.

yl−prolyl cis−trans isomerization is "catalysis by distortion"; that is, the enzyme stabilizes a transition state with an "out-of-conjugation" amide group.[427−430] For the tetrapeptide ALPF the uncatalyzed rate constant for cis−trans isomerization (k_{non}) is 9.7×10^{-3} s^{-1}. The k_{cat}/K_m for the FKBP12 catalyzed reaction is 2.2×10^6 M^{-1} s^{-1}.[428] This translates into a catalytic proficiency for FKBP12 of only 2.3×10^8 M^{-1}. The out-of-conjugation carbonyl of FK506 has been suggested to act as a "twisted amide" transition state surrogate (see Figure IM*4) for the cis−trans isomerase reaction catalyzed by FKBP12.[427,428]

Since FK506 binds to FKBP12 without having immunosuppressant consequences and the resulting FK506-FKB12 complex binds to and inhibits a second protein calcineurin, it is often referred to as a dual domain compound.[415] The portions of FK506 which contact FKBP12 (enclosed in box in Figure IM*5) are

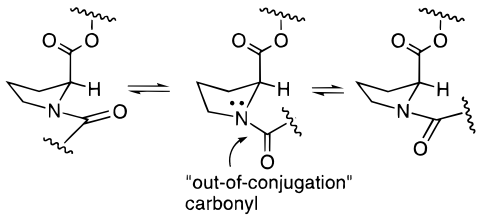
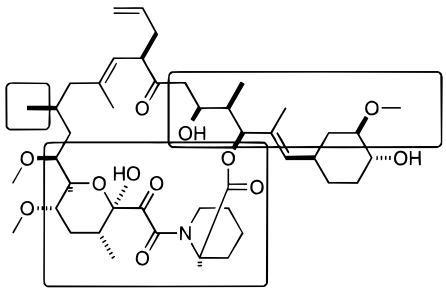


Figure IM*4. The proposed mechanism for PPIase activity of FKBP12 is "catalysis by distortion". The enzyme has evolved to stabilize the an "out-of-conjugation" amide carbonyl.

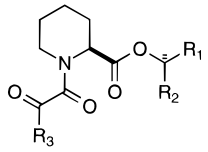
referred to as the "binding domain", while the solvent exposed regions which contact calcineurin are often referred to as the "effector domain".

Rapamycin makes similar contacts with FKBP12 as does FK506.[423] This is not unexpected since they have similar structures of their "binding domains".



FK506

Figure IM*5. FK506 "binding domain" (shown in boxes): ligand atoms within 4 Å of FKBP12.



		K_i app (nM)
IM-2	R_1 = Me, R_2 = H, R_3 = c-Hexyl	2000.
IM-3	R_1 = Me, R_2 = H, R_3 = CMe_2Et	660.
IM-4	R_1 = Ph, R_2 = H, R_3 = CMe_2Et	186.
IM-5	R_1 = H, R_2 = CH_2CH_2Ph, R_3 = CMe_2Et	110.
IM-6	R_1 = Ph, R_2 = CH_2CH_2Ph, R_3 = CMe_2Et	7.

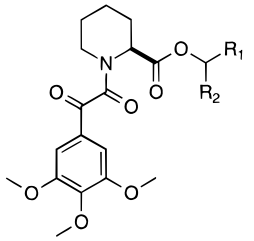
Figure IM*6. Structures and FKBP12 affinities of various ligands.

The C24 (PDB entry: 1FKB numbering) hydroxyl group of rapamycin makes a hydrogen bond (2.70 Å) with the backbone carbonyl of Glu-54. Due to the single bond between C35 and C36, the cyclohexyl group makes more extensive protein−ligand surface contacts. In addition, the C40 hydroxyl group makes a hydrogen bond with the backbone carbonyl of Gln-53. The major structural difference between these two complexes is the size and shape of the solvent-exposed regions. That is, they have different "effector domains". The FKBP12-FK506 complex binds to calcineurin; the FKBP12-rapamycin complex binds to FRAP.

C. Unnatural FKBP12 Ligands

1. Binding Domains

The SmithKline Beecham group was the first to publish that simple acyclic analogs of the FK506 binding domain are high-affinity ligands for FKBP12.[431−433] Compound **IM-2**, a simple, acyclic, FK506 analog has a $K_{i,app}$ of 2 μM vs FKBP12. Substitution of the cyclohexyl R3 group, with a (dimethylethyl)methyl group resulted in a 3-fold increase in affinity for FKBP12. (See Figure IM*6.) Substitution of an R1 hydrogen with either a phenyl group (**IM-4**) or a phenethyl group (**IM-5**) resulted in a 11-fold and 18-fold increase in FKBP12 affinity, respectively. Compound **IM-6**, incorporating both an R1 phenyl group and an R2 phenethyl group is a very high affinity (7 nM) FKBP12 ligand. A crystal structure for **IM-6** bound to FKBP12 has been reported (entry 1FKG).[434] This complex shows that the pipecolinyl ring and dicarbonyl groups bind identically as they do in FK506. The R3 (dimethylethyl)methyl group fills the same site that the pyran



		Ki (app) (nM)
IM-7	R_1 = CH_2Ph, R_2 = H	110.
IM-8	R_1 = $(CH_2)_2Ph$, R_2 = H	19.
IM-9	R_1 = $(CH_2)_3Ph$, R_2 = H	70.
IM-10	R_1 = $(CH_2)_4Ph$, R_2 = H	100.
IM-11	R_1 = $(CH_2)_3Ph$, R_2 = $(CH_2)_3$-3-Pyr	0.5

Figure IM*7. Structures and FKBP12 affinities of various ligands.

ring of FK506 occupies. The phenethyl R2 group makes hydrophobic surface contacts in the site that the cyclohexyl group of FK506 occupies. The R1 phenyl group makes VDW contacts with Phe-46 as well as VDW contacts with the ethyl group the R3 group. While solved in different space groups the structure of the protein is nearly identical between the complexes of FK506 and **IM-6**. A slight conformation change in the side chain of Phe-46 of FKBP12 takes place to maximize VDW contacts with the ligand. The side chains of Asp-37 are in slightly different conformations in the two complexes. In the FK506 complex it approaches the ligand more closely to make a hydrogen bond with the hemiketal hydroxyl group. The flexible region of the protein between Ala-84 and His-94 is very similar in both complexes. The Agouron group has solved the structure of the complex between **IM-4** and FKBP12. Compound **IM-4** binds identically to **IM-6** with the exception that a hydrogen replaces the R2 phenethyl group. The solved structures of the **IM-4** and **IM-6** complexes and the 11-fold increased potency of **IM-6** shows that the placement of the phenylethyl group along the surface of the protein, in the shallow groove occupied by the cyclohexyl ring in FK506, is worth 1.4 kcal mol^{-1}.

The Vertex group has reported that a 3,4,5-trimethoxy phenyl group is a useful pyran replacement in acyclic FKBP12 ligands.[435] As seen in Figure IM*7, some of these compounds are quite potent FKBP12 ligands. The Vertex group has reported the cocrystal structure of **IM-11** ($K_{i,app}$ = 0.5 nM) with FKBP12 (Figure IM*7). This compound is nearly equipotent to FK506 as a isomerase inhibitor. The crystal structure of **IM-11** with FKBP12 has been solved in space group R3. The structure of the complex shows that the pipecolinyl ring and dicarbonyl groups bind nearly identically as they do in FK506. The trimethoxyphenyl group is not in the plane of the pyran group of FK506, it is partially in conjugation with the ketone carbonyl. This substituent occupies a volume that is not filled in the FKBP12-FK506 complex. In addition, this substituent distorts the flexible loop between Ala-84 and His-94. The phenyl ring occupies roughly the same volume that the cyclohexyl ring of FK506 occupies.

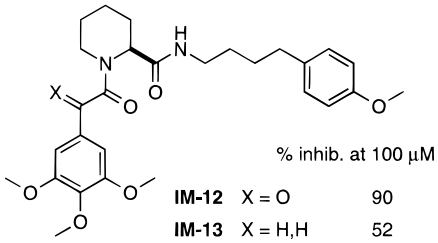


Figure IM*8. Structures and FKBP12 affinities of various ligands.

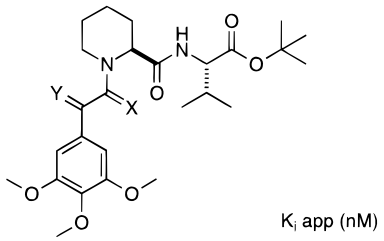


Figure IM*9. Structures and FKBP12 affinities of various ligands.

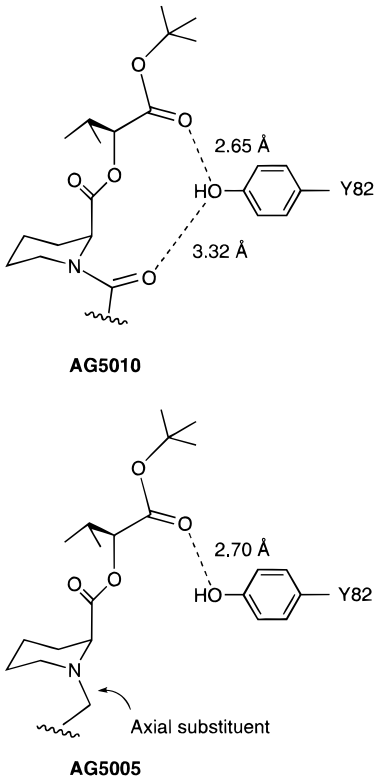


Figure IM*10. Interactions of **AG5010** and **AG5005** with FKBP12.

(A structure of the same complex solved in the p32 space group shows an alternative location for this phenyl group. In this space group the phenyl ring occupies roughly the same space as the cyclohexyl ring of rapamycin.[436]) The other aromatic group, the 3-pyridyl ring, binds in the vicinity of Phe-46 and Tyr-26. In addition, it makes a face to face stacking interaction with the trimethoxyphenyl group.

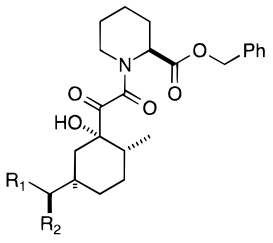
The Abbott group examined the effect of modifications to the ketone carbonyl group in simple acyclic compounds[437] (Figure IM*8). Their results show that ketone **IM-12** is the most potent (IC_{50} = 16 μM) FKBP12 ligand and that changes to that group reduce the affinity of the ligand for FKBP12.

Structural studies by the Agouron group have looked at compounds **AG5010**, **AG5005** and **AG5011** (Figure IM*9) which differ in the tricarbonyl region.[438] These compounds also have peptidyl groups and the structure of their complexes may be used as models for the binding of substrates for the isomerase reaction. (See Figure IM*10.) The structure of the complex of the more potent dicarbonyl compound **AG5010** reveals an slight conformational change in the side chain of Tyr-82. This conformational change places the tyrosine hydroxyl 3.2−3.3 Å from the amide carbonyl oxygen of the ligand and 2.65 Å from the ester carbonyl of the ligand. Thus, it appears that in this complex the hydrogen bond between Tyr-82 and the amide carbonyl of the dicarbonyl moiety has been broken and replaced by a hydrogen bond with the ester carbonyl of the ligand. One of the methyls of the isopropyl group occupies a similar place in space as does the C25 methyl of FK506, making VDW contacts with Phe-46 of FKBP12. The other methyl group makes VDW contacts with the trimethoxy phenyl group of the ligand. The *tert*-butyl group makes VDW contacts with the protein in the same region of space that the cyclohexyl group of rapamycin does.

Considering the hydrogen-bonding arrangement of amide **AG5010**, in which there is no hydrogen bond between Tyr-82 and the C8 amide carbonyl, it is somewhat surprising that the amine **AG5005** has 50-fold less affinity for FKBP12. The structure of the complex revealed a surprise. (See Figure IM*10.) Like **FK506** and **AG5010**, **AG5005** places the ketone carbonyl in the small pocket created by three aromatic CH groups. In the amine **AG5005**, this situation is accomplished by having an axial substituent off the sp3 nitrogen. Thus, the poorer affinity of **AG5005** is likely due to ligand strain; the piperidine ring has two axial substituents.

Substitution of the ketone carbonyl group for an exocyclic methylene results in 25-fold less affinity for FKBP12. The structures of these two complexes are essentially superimposable. It thus appears that the increased affinity for the ketone is due to a more favorable interaction made by the carbonyl group with the small pocket created by three aromatic CH groups than is made by the exocyclic methylene.

The Agouron group has reported that substituted cyclohexanes, derived from (*R*)-(−)-carvone, are useful pyran replacements in acyclic FKBP12 ligands.[439] As seen in the Figure IM*11 ligand **IM-17** is a very potent FKBP12 inhibitor. A crystal structure between **IM-16** and FKBP12 was reported. The pipecolinic ring and dicarbonyl groups bind identically as those groups bind in FK506. The cyclohexyl ring of **IM-16** binds similarly to the pyran ring of FK506. A small conformational change was observed in which the orientation of the cyclohexyl ring, relative to the pipecolinic ring, was different than seen for the pyran



		K_i app (nM)
IM-16	$R_1 = R_2 = Me$	210.
IM-17	$R_1 = CH_2\text{-}iPr$ $R_2 = OMe$	2.8

Figure IM*11. Structures and FKBP12 affinities of various ligands.

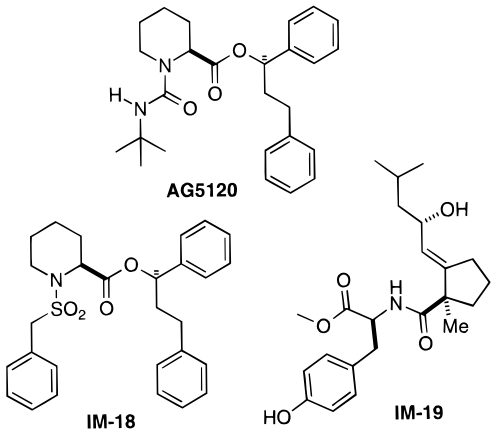
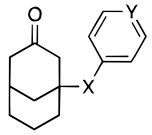


Figure IM*12. Structures of various ligands.

ring in FK506. The observed conformational change allows for maximization of intramolecular hydrophobic interactions between the benzyl ester and the cyclohexyl groups.

The Agouron group has reported SAR studies and several cocrystal structures of urea-containing FKBP12 ligands.[440] One of the more potent compounds reported is **AG5120** ($K_{i,app}$ = 350 nM). (See Figure IM*12.) This compound shows a 15-fold loss in affinity for FKBP12 compared to the related ketoamide **IM-6**. During the design process, it was anticipated that the ureas would forfeit the beneficial interactions between the ketone carbonyl and the small pocket created by three aromatic CH groups. However, it was expected that the urea NH would gain a new interaction by forming a hydrogen bond to the carboxylate of Asp-37 analogous to that made by the hemiketal hydroxyl of FK506. The crystal structure of the complexes between **AG5120** and **IM-6** with FKBP12 were solved in the same space group ($P2_12_12_1$). Comparison of these two complexes show that the common features of these two ligands (pipecolinyl ring, ester, and 1,3-diphenylpropane) bind almost identically. Thus, the pipecolinyl ester is functioning as an anchor. The urea functionality binds as expected. The urea carbonyl forms a hydrogen bond to the phenol of Tyr-82 (O−O distance: 2.7 Å) and the urea NH forms a hydrogen bond to one of the anti-lone pairs of Asp-37 carboxylate (O−N distance: 3.15 Å). To accommodate the formation of this hydrogen bond, a slight conformational change in the protein around Asp-37 occurs. The absence of the ketone carbonyl results in the small pocket created by the three aromatic CH groups being left vacant. No protein conformational change takes



(±)

		K_i app (μM)
IM-20	X = O, Y = CH	116.
IM-21	X = S, Y = CH	9.2
IM-22	X = SO_2, Y = CH	48.5
IM-23	X = S, Y = N	7.9

Figure IM*13. Structures and FKBP12 affinities of various ligands.

place to remove that pocket. The presence of the unfilled pocket results in a gap between the protein and the urea ligand. The 15-fold loss of FKBP12 affinity of the urea **AG5020** relative to the keto amide **IM-6** and the observed structures of their complexes suggest that the creation of the gap between the protein and the ligand is energetically costly, and it is not offset by the formation of the hydrogen bond between the urea NH and Asp-37.

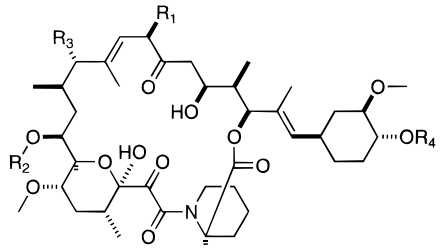
The SmithKline Beecham group has reported on extensive SAR work on simplified acyclic binding cores for FKBP12 ligands. Their data suggests that the keto amide functionality is superior to many other functionalities. Another group which imparts significant affinity to ligands is a sulfonamide. Ligand **IM-18** was reported to have $K_{i,app}$ = 160 nM against FKBP12. No structural data for sulfonamide-containing ligands has appeared in the literature.

The Virginia and Wyest-Ayest groups have reported on novel, mechanism-based inhibitors of the isomerase activity of FKBP12.[441] Compound **IM-19**, prepared as an inseparable mixture of two stereoisomers, has a K_i = 8.6 μM as an inhibitor of FKBP12. No structural data has been reported for this compound or related ligands.

The Agouron group reported the design and evaluation of a novel class of FKBP12 ligands[442] (Figure IM*13). The design process was assisted by the use of Ludi,[443] a suggestive program for *de novo* ligand design. Ligand **IM-21** had a $K_{i,app}$ = 9.2 μM against FKBP12. The thioether was 1 order of magnitude more potent than ether **IM-20**, and 5-fold more potent than sulfone **IM-22**. A cocrystal structure between **IM-23** and FKBP12 was described. In this complex the bicyclic ring system binds in the site occupied by the pipecolinic ring of FK506 and makes VDW contacts with Ile-56, Val-55, Trp-59, Phe-46, Phe-99, and Tyr-26. The aromatic ring binds in the site occupied by the pyran ring of FK506, and the carbonyl group forms a hydrogen bond to the backbone NH of Ile-56 (O−O distance: 2.8 Å). The sulfur atom was observed to fill the pocket lined by three aromatic ϵ-hydrogens, one each from Tyr-26, Phe-36, and Phe-99 (S−C distances: Tyr-26 4.2 Å, Phe-99 3.8 Å, Phe-36 4.1 Å). This is the same pocket filled by the ketone carbonyl of FK506.

2. Effector Domains: FK506 Analogs

Much work has been done on exploring modifications to the effector domain of FK506 (Figure IM*14). A brief summary follows. Ascomycin (**IM-24**) and



	R_1	R_2	R_3	R_4
FK506	$CH_2CH{=}CH_2$	CH_3	H	H
Ascomycin (**IM-24**)	CH_2CH_3	CH_3	H	H
FK523	CH_3	CH_3	H	H
IM-25	$(CH_2)_3CH_3$	CH_3	H	H
IM-26	CH_2CH_3	H	H	H
L-685,818	CH_2CH_3	CH_3	OH	H
IM-27	CH_2CH_3	CH_3	H	CH_2Ph

Figure IM*14. Structures of various ligands.

FK523 are natural products closely related to FK506. All three compounds have nearly the same affinity for FKBP12. However, their FKBP12 complexes have different effects on calcineurin (FK506, $K_i = 34$ nM, ascomycin, $K_i = 89$ nM; FK523, $K_i = 230$ nM).[421] Semisynthetic modifications on FK506 produced **IM-25**, in which the allyl group has been exchanged for an *n*-butyl group. Compound **IM-25** retains high affinity for FKBP12; however, **IM-25** is not immunosuppressive (it presumably does not inhibit calcineurin).[444] A demethyl analog of ascomycin, **IM-26**, retains significant FKBP12 affinity ($K_{i,app} = 15$ nM); however, it is a poor inhibitor of calcineurin ($K_i = 1.6\ \mu$M).[421] Another semisynthetic analog of ascomycin, **L-685,818** binds tightly for FKBP12; however, it has no effect on calcineurin inhibition.[445] These data suggest that calcineurin inhibition is highly specific.

The structure of the **L-685,818**/FKBP12 complex (PDB entry: 2FKE) has been determined and is nearly identical to the FK506 complex. Since the conformations of FK506 and **L-685,818** are nearly identical when bound to FKBP12, it was proposed that their different biological profile was due to the presence of the additional hydroxyl group of **L-685,818** which is solvent exposed in the complex.[446]

In contrast to the above examples, modifications to the C32 hydroxyl group of the cyclohexyl group (i.e., **IM-27**) did not have a significant effect on immunosuppressive activity and a wide variety of groups were tolerated.[444]

3. Effector Domains: Unnatural FKBP Ligands and Hydrophobic Collapse

An acyclic analog of FK506, **SBL506** (Figure IM*15), has been described by a Harvard group.[447] This compound shows greatly reduced affinity for FKBP12 ($K_i = 207$ nM). However, the **SBL506**/FKBP12 complex was reported to be a potent inhibitor of calcineurin ($K_i = 330$ nM). The Fisons group have described the related ligands **IM-28** and **IM-29**.[448] Like **SBL506** both compounds show reduced affinity for FKBP12 compared to FK506 (**IM-28** K_d = 140 nM, **IM-29** $K_d = 300$ nM). Unlike **SBL506**, neither **IM-28** nor **IM-29**, in complex with FKBP12, inhibit calcineurin. The discrepancy in calcineurin activity between **SBL506** and compounds **IM-28** and **IM-29** is quite interesting. This suggests that the calcineurin activity of **SBL506** is highly specific. Another acyclic FK506 analog **IM-30** has been reported by the Fisons group.[449] This compound shows reduced affinity for FKBP12 ($K_d = 400$ nM). The FKBP12/**IM-30** complex does not inhibit calcineurin. No structural data for these interesting compounds have been reported.

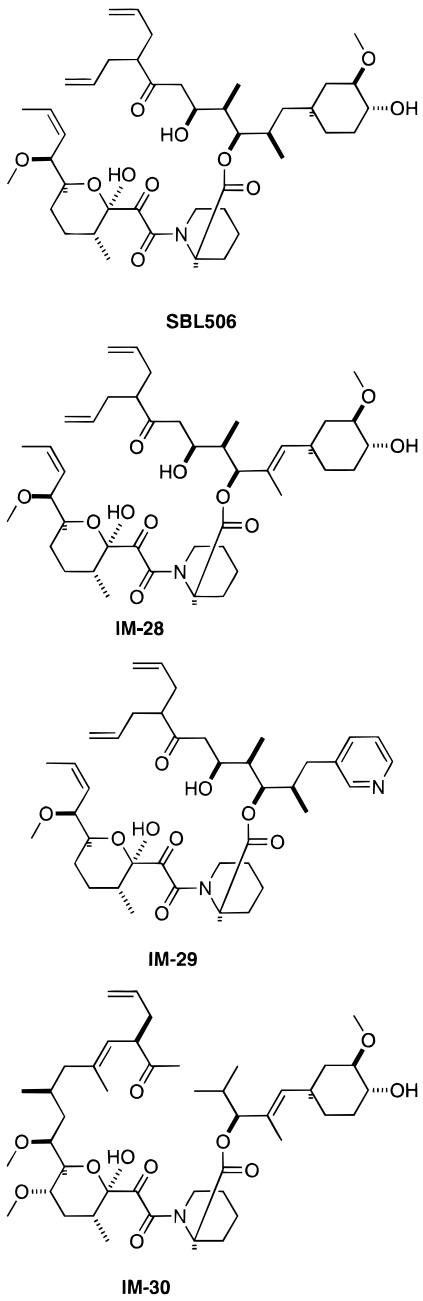


Figure IM*15. Structures of various ligands.

The Agouron group has prepared a number of simplified acyclic FKBP12 ligands.[450] The 2-ethylbutyric ester **AG5163** (Figure IM*16) was found not to bind to FKBP12. The parent alcohol **AG5149** was found to have relatively high affinity for FKBP12 ($K_{i,app} = 58$ nM). Affinity for FKBP12 decreased as the size of the ester group was increased (**AG5154**,

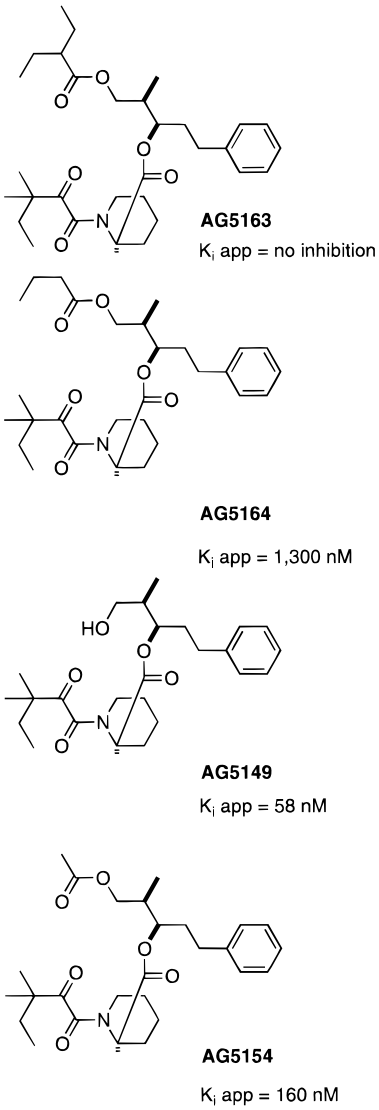


Figure IM*16. Structures and FKBP12 affinities of various ligands.

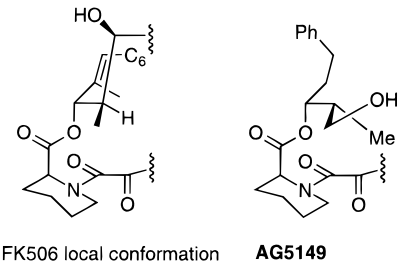


Figure IM*17. The side chain of **AG5149** binds to FKBP12 with a different conformation than **FK506**.

$K_{i,app}$ = 160 nM; **AG5164**, $K_{i,app}$ = 1300 nM). High-resolution crystal structures for alcohol **AG5149** and ester **AG5164** complexed with FKBP12 have been obtained.

The structure of the **AG5149**/FKBP12 complex showed a binding mode similar to what was expected. (See Figure IM*17.) The dimethylethyl and phenylethyl groups bound as observed for **AG5122**. The hydroxy ethyl side chain, however, bound differently than expected on the basis of the structure of FK506. Whereas the hydroxyl group of FK506 forms a hydrogen bond to the backbone carbonyl of Glu-54, the hydroxyl methyl group of **AG5149** makes VDW

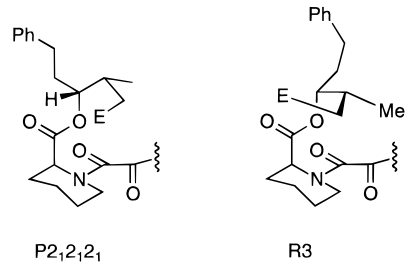


Figure IM*18. The bound conformations of **AG5164** (E = *n*-Pr-O_2C) in two different space groups show that the E side chain has undergone hydrophobic collapse with the surface of the protein.

contacts with the aromatic side chain of Phe-46. The hydroxyl group makes no strong protein−ligand contacts and is somewhat solvent exposed. The distinct binding modes for the hydroxyl groups in FK506 and **AG5149** are achieved by having different rotomers about C25-C26 (FK506 numbering).

The structure of the much less potent FKBP12 ligand **AG5163** was solved in two different space groups ($P2_12_12_1$ and $R3$). (See Figure IM*18.) The $P2_12_12_1$ structure showed an unexpected binding mode for the butyryl side chain. This group makes VDW contacts with several surface residues (Phe-46, Phe-48,Glu-54, Val-55, and Lys-47) on FKBP12. This group does not bind in a well-defined cleft, but rather lies along the protein surface. This is accomplished by having a different rotomer about C25−C26 than either FK506 or **AG5149**. The dimethylethyl and phenylethyl groups bound as observed for **AG5122** and **AG5149**. The structure solved in the $R3$ space group shows a different bound conformer. The phenylethyl group is bound differently, presumably due to crystal packing. While the butyryl side chain is not defined in the electron density, the location of C24, C25, and C26 require that this side chain lie along the protein surface in a similar location as seen for the $P2_12_12_1$ structure. This is accomplished by having a rotomer about C25−C26 similar to **AG5149**. Thus, these two structures have different bound conformations of their side chains. What is apparent from the two structures of **AG5164** is that the butyryl group has undergone hydrophobic collapse[125] onto the surface of the protein.

In contrast to FK506 which exposes a hydrophobic effector domain into solvent when complexed to FKBP12, the hydrophobic side chain of **AG5164** makes VDW contacts with the protein surface. The bound structures observed for **AG5164** demonstrate that a conformation mimicking FK506 and exposing a hydrophobic surface into solvent is not favorable. Alternative conformations in which the flexible hydrophobic group undergoes hydrophobic collapse onto the protein surface are experimentally observed and presumably more stable. The lack of FKBP12 affinity observed for **AG5163** can be rationalized on the basis of the experimental structures of **AG5164**. The larger branched ester group in **AG5163** cannot be accommodated on the surface of the protein, and an FK506-like conformation exposing hydrophobic surface into solvent is not energetically favorable. In addition, the low FKBP12 affinities of **AG6163** and **AG5164** may also be due to unproductive hydrophobic collapse in the solution conformations of the

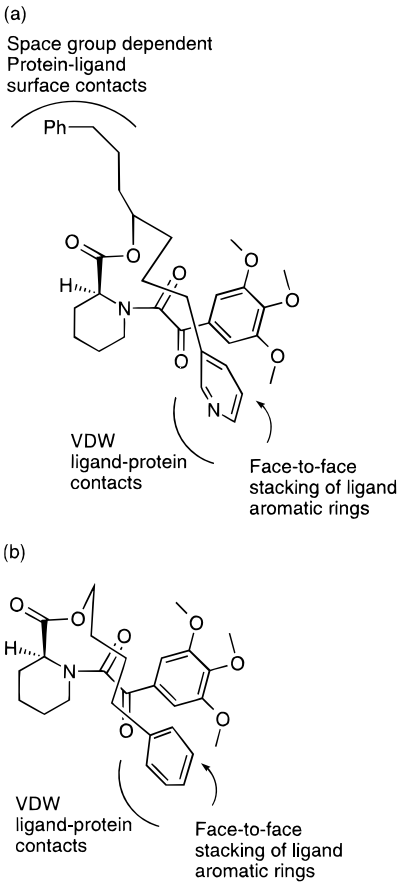
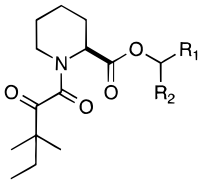


Figure IM*19. (a) Ligand **IM-11** has two $(CH_2)_3Ar$ side chains: one side chain makes close VDW contacts with the trimethoxyphenyl ring of the ligand and weak VDW contacts with the protein, and the other side chain makes extensive VDW contacts with the protein surface. (b) Ligand **IM-9** has only one $(CH_2)_3Ph$ side chain; the side chain makes close VDW contacts with the trimethoxyphenyl ring of the ligand and weak VDW contacts with the protein.

ligands or due to many ligand conformations in solution. Many of these solution conformations may not be preorganized to bind efficiently to FKBP12.

The hydrophobic collapse between the ligand and the protein observed for **AG5164** is not an isolated case. The Agouron group has observed this phenomenon in a number of FKBP12 ligands.[451]

The previously discussed ligand **IM-11** has two nearly identical $(CH_2)_3Ar$ side chains. The solved structure[435] shows that one side chain places a phenyl ring in the shallow grove occupied by the cyclohexyl ring of FK506 (another structure in a different space group shows an alternative location for this phenyl group)[436] and the other side chain places the pyridyl ring in the vicinity of Phe-46 and Tyr-26. In addition, this pyridyl ring makes a face to face stacking interaction with the trimethoxyphenyl group of the ligand. Thus, one arm makes extensive VDW contacts with the protein, while the other arm makes less contact with the protein but makes extensive VDW contacts with another part of the ligand. (See Figure IM*19.) Ligand **IM-9** contains only a single $(CH_2)_3Ph$ side chain. The structure of the FKBP12/**IM-9** complex shows that this side chain makes little contact with the protein but makes a face to face stacking interaction with the trimethoxyphenyl group



		K_i app (nM)
AG5387	$R_1 = (CH_2)_3Ph$, $R_2 = H$	326.
AG5390	$R_1 = (CH_2)_3Ph$, $R_2 = (CH_2)_3Ph$	NI

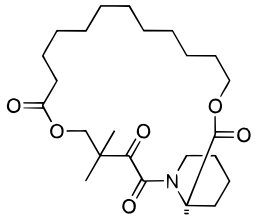
Figure IM*20. Structures and FKBP12 affinities of various ligands.

of the ligand.[436] In the case of **IM-9** the observed complex shows that the ligand has undergone hydrophobic collapse onto itself. For **IM-11**, with an additional side chain, one side chain undergoes hydrophobic collapse onto the trimethoxyphenyl ring while the other side chain undergoes hydrophobic collapse onto the surface of the protein. On the basis of the relative affinities of **IM-9** and **IM-11** for FKBP12 the VDW contacts of the $(CH_2)_3Ph$ group of **IM-11** with the protein surface is worth about 2.9 kcal mol^{-1}.

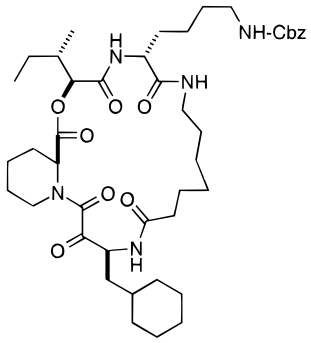
Two ligands related to **IM-9** and **IM-11** have been prepared (Figure IM*20). Compound **AG5387**, having one $(CH_2)_3Ph$ side chain, is a potent FKBP12 ligand having $K_{i,app} = 326$ nM. Surprisingly, **AG5390**, having two $(CH_2)_3Ph$ side chains, has no measurable affinity for FKBP12. No structure for the **AG5387**/FKBP12 complex is available. However, on the basis of the activities of **AG5387** and the structurally related **AG5149** it is possible that the side chain of **AG5387** binds along the protein in the shallow groove that is occupied by the phenyl group of **IM-11**. The presence of the bulkier Me_2CEt side chain of **AG5390** presumably does not allow the second $(CH_2)_3Ph$ side chain to bind in a similar manner to **IM-9** or **IM-11**; thus, **AG5390** suffers a fate similar to **AG5163**. The second $(CH_2)_3Ph$ group is too large to be accommodated on the surface of the protein, and a conformation exposing hydrophobic surface into solvent is not energetically favorable.

On the basis of the huge difference in FKBP12 affinities between **IM-11** and **AG5390**, it is tempting to speculate that the ligand−ligand hydrophobic collapse seen in the complexes of **IM-9** and **IM-11** may also occur in aqueous solution. This productive hydrophobic collapse would serve to preorganize ligands **IM-9** and **IM-11** for binding to FKBP12. In addition, minimizing exposed hydrophobic surface for **AG5390** (unproductive ligand hydrophobic collapse) in aqueous solution may result in conformers which are not complimentary to FKBP12.

These structures clearly illustrate an important feature of protein−ligand interactions. It is well known that one of the driving forces in protein folding is internal packing of hydrophobic groups and the minimization of exposed hydrophobic surface area.[452] It is also known that burying ligand hydrophobic surface area in hydrophobic sites on a protein provides an important driving force in protein−ligand complex formation.[452] Many acyclic FKBP ligands attempt to expose hydrophobic surface into solvent; however, the available data suggests that these



IM-31



IM-32

Figure IM*21. Structures of unnatural macrocyclic FKBP12 ligands.

ligands bind to FKBP12 in a conformation which minimizes solvent exposed hydrophobic surface area. These new examples again demonstrate the importance of hydrophobic collapse in molecular recognition processes in aqueous solution.

FK506 and rapamycin, in part because of their macrocyclic nature, are able to avoid protein−ligand hydrophobic collapse and expose significant hydrophobic surface area into solvent. Structures of other macrocyclic ligands in complex with FKBP12 have been described.

A cyclic peptide-FK506 hybrid ligand **IM-32** (Figure IM*21) has been described.[453] It has $K_{i,app} = 210$ nM vs FKBP12. As observed for FK506, a strong hydrogen bond between the axial carbonyl of the six-membered ring and the backbone NH of Ile-56. For the carbonyl group of the inhibitor's isoleucine, a hydrogen bond was observed with the phenolic hydroxyl of Tyr-82. The amide NH of the lysine of the inhibitor forms a hydrogen bond with the backbone carbonyl of Glu-54. These interactions place the isoleucine side chain of **IM-32** at the same location of the C24 and C25 ring atoms, and C25 methyl group of FK506. Thus, the dipeptide fragment of **IM-32** binds to FKBP12 by forming a pair of hydrogen bonds characteristic of antiparallel β-sheets. (See Figure IM*22.) The hydrogen-bond scheme observed was considered as a model for binding of a protein (or peptide substrate).

A very similar hydrogen-bonding scheme was observed for **AG5010** and related compounds, as discussed above **AG5010** is a *tert*-butyl ester and cannot form a hydrogen bond with the backbone carbonyl of Glu-54. Comparison of the structures of **IM-32** and **AG5010** shows that this hydrogen bond formation occurs by FKBP12 having undergone a small conformation change near Glu-54. Compound **AG5089**, an analog of **AG5010** having a *tert*-butyl amide instead of a *tert*-butyl ester, was prepared and found to be much less potent as an FKBP12 ligand ($K_{i,app}$ = 2,800 nM).[454] Thus, the contribution of this hydrogen bond to the overall strength of these complexes is not clear.

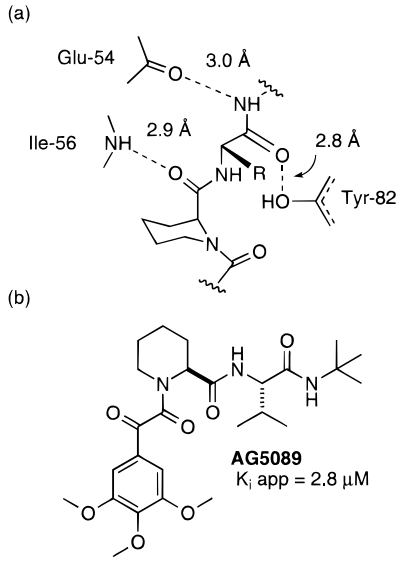
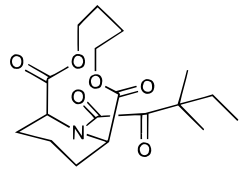


Figure IM*22. (a) Interactions of peptide side chain with FKBP12. (b) *tert*-Butyl amide **AG5089** is a weaker FKBP12 ligand than the *tert*-butyl ester **AG5010**.

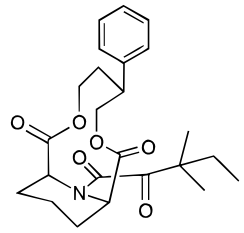
The $(CH_2)_6$ linking group of the macrocycle of **IM-32** follows a much different path than does the macrocycle of FK506 or **IM-31**. This $(CH_2)_6$ linking group of the macrocycle causes a major conformational change in the protein for His-87 to Ile-90. The linking side chain makes VDW contacts with the altered surface of the protein. Superimposing the FKBP12 complexes of **IM-32** and FK506 shows that the linking chain of **IM-32** would have a major steric interaction with His-87 if this protein conformational change did not take place. This $(CH_2)_6$ hydrophobic linking group is, however, quite solvent exposed.

A 21-membered macrocyclic ring ligand (**IM-31**) related to FK506 has been described.[434] It has a $K_{i,app}$ = 100 nM vs FKBP12. The macrocyclic ring contains 11 methylene units as opposed to the highly functionalized nature of the FK506 ring. Its structure with FKBP12 shows that the macrocyclic ring does not project into solvent, but rather loops around the pipecolinate binding pocket in the same direction as the ring of FK506, and makes VDW contacts with the sides chains of Val-55, Phe-46 and the Arg-42/Asp-37 salt bridge. In this example, the macrocyclic ring has undergone hydrophobic collapse with the surface of the protein.

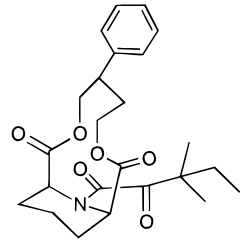
The Agouron group has reported on [7.3.1] and [8.3.1] macrocyclic ligands in which C2 and C6 (FK506 numbering) of a pipecolinic ring are tied back into a ring.[455] A crystal structure of the complex between **AG5230** ($K_{i,app} = 10\ \mu$M) and FKBP12 shows that the bridging four carbon chain is solvent exposed. A substituted analog **AG5397** was prepared as an inseparable 1:1 mixture of two diastereomers (**AG5397_A** and **AG5397_B**, Figure IM*23).[456] The $K_{i,app}$ of this mixture was 1.1 μM. This mixture was cocrystalized with FKBP-12. The structure, solved in space group R3, showed that only isomer **AG5397_B** was bound to the protein. In this case,



AG5230



AG5397_A



AG5397_B

Figure IM*23. Macrocyclic FKBP12 ligands. The structure of the complex between FKBP12 and **AG5397_B** shows that the phenyl ring is solvent exposed.

diastereomer **AG5397_B** presumably has higher affinity for the protein than isomer **AG5397_A**. The structure of **AG5397_B** showed that the phenyl ring was solvent exposed. The [8.3.1] macrocyclic ring system is thus capable of avoiding ligand hydrophobic collapse upon the protein surface.

D. FKBP12-FK506–Calcineurin Complex

The structure of the complex between FK506/FKBP12 and CaNA/CaNB has been solved by both the Vertex[457] and Agouron groups.[458] The FK506/FKBP12 complex binds at a site remote from the active site of calcineurin and the FK506 ligand is located about 25 Å from the phosphatase active site.

The FK506/FKBP12 complex binds at a portion of the CaNA/CaNB interface. Specifically, the CaNB-binding α-helix of CaNA (BBH) and the adjacent regions of CaNB. The principal site of interaction between FK506 and calcineurin is a predominately hydrophobic cleft located at the interface of CaN-B and the BBH of CaNA. Side chains that form the cleft come from residues Leu-343, Pro-344, Trp-352, Ser-353, and Phe-356 of CaNA and residues Leu-115, Met-118, Val-119, and Leu-123 of CaN-B. This binding cleft is about 8 Å long and is complimentary to the surface of C15−C21 of FK506. The majority of contacts made by FK506 are from C15−C17 and the C21 allyl group. The allyl group of FK506 makes intimate VDW contacts in the pocket defined by Leu-116, Met-119, and Val-120 of CaNB and Trp-352 and Phe-356 of CaNA.

The only polar interaction between FK506 and calcineurin is a bifurcated hydrogen bond between the Nϵ1 of Trp-352 and the ether oxygens of the C13 and C15 methoxy groups. The C29−C34 cyclohexyl group of FK506 is partially solvent exposed and does not appear to make tight, specific contacts with calcineurin.

Two regions of FKBP12, Lys-34 to Lys-47 and His-87 to Ile-90, also contact calcineurin in the complex and bury approximately 800 Å^2 of solvent assessable surface area. The side chains of His-87, Pro-88, and Ile-90 make hydrophobic surface contact with part of the BBH. The Lys-34 to Lys-47 regions of FKBP12 do not appear to make tight, specific contacts with calcineurin. The Arg-42 side chain of FKBP12 makes VDW contacts with CaNA and the guanidine group interacts with the phenolic hydroxyl of Tyr-341 of CaNA. The P89G + K90I double mutant of FKBP13 which has a glutamine at residue 42, in complex with FK506, is only 2-fold less potent as a calcineurin inhibitor.[419] This data suggests that the polar interaction made by Arg-42 of FKBP12, with the phenolic hydroxyl of Tyr-341 of CaNA, is not essential for efficient calcineurin inhibition.

The conformation of FK506 in the ternary complex is nearly identical to the conformation seen in the structure of the FK506/FKBP12 complex. The conformation of FKBP12 is also similar in both complexes. However, the relative position of FK506 to FKBP12 in the ternary complex is slightly different than that in the binary complex. FK506 is rotated about 8° from the body of FKBP12. This results in a displacement of 1.7 Å for C21 at the base of the allyl group. The change in position of FK506 results in the loss of the weak hydrogen bond between the carbonyl oxygen of Glu-54 and the C24 hydroxyl of FK506. In addition, a slight conformational change in His-87 to Ile-90 of FKBP12 is also observed. The side chain of Ile-90 adopts a different conformation in the ternary complex. These slight conformational changes allow the FK506/FKBP12 complex to make more intimate contact with calcineurin.

From the structures of the FK506/FKBP12/CaNA/CaNB ternary complex the mechanism of enzyme inhibition is not completely clear. In the ternary complex, no atoms of FKBP12 are closer than 10 Å from the calcineurin phosphatase active site.

One proposal for CaN inhibition suggests that the bound FKBP12 physically blocks access of protein and the polypeptide substrates to the CaN phosphatase active site.[457] Consistent with this proposal, the observed binding of FK506/FKBP12 to CaN shows that portions of FKBP12 bind to a surface groove which is 10−20 Å from the active site. A second proposal suggests that inhibition is caused by subtle changes in active-site geometry caused by binding of the FK506/FKBP12 complex.[458] Consistent with this proposal, FK506/FKBP12 exhibits classical noncompetitive inhibition of CaN with the RII peptide and with the protein substrate, DARPP-32.[458] Noncompetitive inhibition generally implies formation of a catalytically inactive enzyme−substrate−inhibitor complex. Consistent with both proposals is that FK506/FKBP12 binding results in a slight increase in CaN phosphatase activity when *p*-nitrophenyl phosphate is used as the substrate instead of a phosphopeptide. Noncompetitive inhibi-

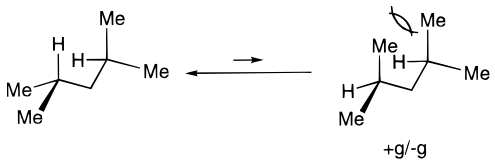


Figure IM*24. 2,4-Dimethylpentane has one preferred conformation which avoids the high energy +gauche/−gauche (+g/−g) pentane interaction. The +g/−g conformation has 1,5 VDW repulsion between two methyl groups.

tion would be consistent with the steric argument if the FK506/FKBP12 complex allows substrate binding but does not allow formation of a productive enzyme−substrate complex.

E. Why Does FK506 Work?

FK506 is a macrocyclic ligand which is able to expose hydrophobic surface area into solvent in its complex with FKBP12. This exposed surface area is then capable of binding tightly to a different protein, calcineurin, without undergoing a major conformational change. In contrast, only one unnatural ligand (excluding semisynthetic FK506 analogs), **SBL506**, has been reported to have this remarkable property. As has been described, many acyclic ligands bind to FKBP12 with a conformation which minimizes the amount of exposed hydrophobic surface area (i.e., they often undergo some form of hydrophobic collapse). This poses the important question, "what structural features of FK506 are responsible for its property as a chemical inducer of dimerization"?

Examination of the structure of FK506 reveals a 21-membered macrocyclic ring with several methyl and methoxy groups plus an allyl group in 1,3 relationships to each other. The crystal structure of the complex between FK506 and FKBP-12 shows that these structural features are acting as local conformational locks. There are now several structures of FK506 which reveal its conformation: the structure of it in complex with FKBP12, its ternary complex with FKBP12 and calcineurin, its small molecule crystal structure (CSD entry: FINWEE10),[404] its NMR structure in $CDCl_3$,[459] and an NMR structure of a water soluble analog in D_2O.[460] The NMR structures of FK506 show that they exist in solution as a 1:1 mixture of cis and trans amide rotomers. The small molecule crystal structure shows only the cis amide rotomer. Structures of the FK506/FKBP12 complexes show that only the trans rotomer binds to the protein. From these different experimental structures it is quite clear that FK506 does not have a rigid structure with one unique low-energy conformation. Rather, it has several experimentally determined conformations for the macrocyclic ring.

Conformations of acyclic segments of molecules, such as the polyether ionophores, can be rigidified by avoidance of relatively high energy +gauche/−gauche pentane interactions (see Figure IM*24).[461] The macrocyclic ring of FK506 has several methyl and methoxy groups plus an allyl group in 1,3 relationships to each other. It is likely that these groups restrict the conformation space available to

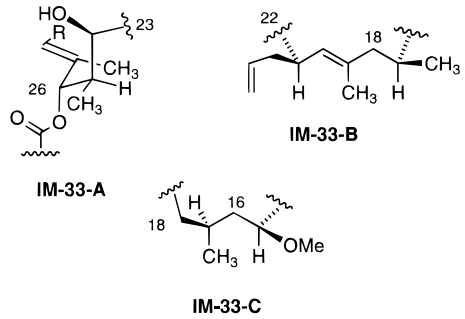


Figure IM*25. Local conformation locks found in the bound conformation of FK506.

the macrocycle by making some conformations high in energy. However, they do not provide a single conformational lock for the macrocyclic ring.

Examination of the experimentally determined conformations of FK506 shows that there are at least three sets of independent conformational locks, the region from C19−C29 (**IM-33-A**), the region from C14−C18 (**IM-33-B**), and the region from C16−C22 (**IM-33-C**). (See Figure IM*25.) In each FK506 conformation there are no severe 1,5 VDW repulsions in the regions of these local conformational locks. For the region between C19−C29, only two different local conformations have been experimentally observed. For FK506 in all of its complexes with proteins, in aqueous solution (both cis and trans rotomers), in its single-crystal structure (cis rotomer) and as its trans rotomer in $CDCl_3$ solution a single conformation (**IM-33-A**) for this region has been observed. In the FK506/FKBP12 complex this local conformation results in the introduction of key hydrophobic regions of FK506 into solvent; these regions make close contacts with calcineurin in the ternary complex. For the region from C16 to C22, three distinct local conformations have been observed which differ by rotation around the C18−C19 bond. In all experimentally determined conformations of FK506, a single local conformation is observed for C18−C22; this is due to avoidance of $A^{1,3}$ strain with the allylic C17 methyl group. For the region between C14 and C18, five distinct local conformations are been observed in six experimental structures. The conformation of FK506 bound to a mutant FKBP has yet another conformation.[462] The presence of a methyl and methoxy group in a 1,3-relationship to each other does not appear to be a strong conformational lock.[463] In the bound conformation of FK506 the C18 carbon and the ether oxygen of the C15 methoxy group are 3.32 Å apart.

Each of the local conformations **IM-33-A-C** have in common no serious VDW repulsion. However, the various experimental structures reveal a conformationally heterogeneous nature for FK506. It appears likely that the macrocyclic ring, working in concert with the ring substituents, allows for only a few energetically favorable conformations in a given medium.

NMR relaxation methods have shown that the solvent exposed portions of FK506 are quite rigid in the complex with FKBP12.[464,465] These studies show that the macrocycle backbone, especially the C21 to C25 region, is relatively immobile on the picosecond time scale. The regions around C14 and C15 are

more mobile. The solvent exposed allyl group appears to be disordered.

The R42I single mutant of FKBP12 in complex with FK506 shows a 180-fold loss of activity as a calcineurin inhibitor. The structure of FK506 bound to the single mutant R42I of FKBP12 has been determined.[466] The structure of FK506 in the mutant complex is nearly identical with that of the wild type. No conformation changes in either the 40s or the 80s loop of the mutant FKBP12 are observed. In addition, the B factors for FK506 are not higher than that of the wild type. In this R42I mutant complex structure, two tightly bound water molecules substitute for the missing arginine guanidino nitrogens of R42 and are well accommodated in the gap created by the smaller isoleucine substituent.

On the basis of the structure of the FK506/FKBP12/CaN complex it is possible that the decreased calcineurin inhibition is due to unfavorable VDW contacts by the branched chain of Ile-42 with CaN, and/or burying two ordered water molecules which substitute for the missing arginine guanidino nitrogens of R42.

The H87V and R42K single mutants of FKBP-12, and the R42K+H87V double-mutant of FKBP-12 in complex with FK506 exhibit a 4-fold, 180-fold, and 680-fold reduction in CaN inhibitory activity respectively. Structural studies for these mutant FKBP-12/FK506 complexes have also been reported.[462] A crystal structure of the complex between FK506 and the R42K+H87V double mutant (PDB entry: 1BKF) revealed a significant conformational change for the bound FK506. In this complex the macrocyclic backbone adopts an alternative conformation in which there is a 3.5 Å shift for C17 and as much as a 5.5 Å shift for the C17-methyl group. In addition, significant changes were seen in the position of the C15-OMe and the C19 Me groups. For both single mutants, the conformation of FK506 closely resembles that seen in the native complex structure. However, differences in ligand temperature factors were observed. The H87V and R42K+H87V mutants show higher ligand temperature factors while the R42K mutant shows only slightly higher temperature factors than the native complex. These higher temperature factors are consistent with higher ligand mobility. Higher temperature factors for the 40s loop of the R42K mutant are also observed. (The resulting destabilization of the 40s loop is observed even though the conformation of mutant protein complex still closely resembles that of the native and FK506 binding and PPIase activity are preserved.) This information suggests that the rigidity of the native FKBP-12/FK506 complex is due to both features of the ligand and its interactions with the protein. These experiments show that the conformation of the C14−C19 region of FK506 is also dependent upon the precise nature of the protein.

The macrocyclic ring of FK506, with its various substituents acting as local conformational locks and in concert with binding to FKBP-12, generates a rigid ligand in the complex. While these ring substituents alone are not capable of enforcing a single conformation, in the presence of FKBP12 FK506 appears to become conformationally homogeneous. Since the conformation of FK506 is nearly identical in its binary and ternary complexes, it appears likely that an important role played by FKBP12 in inhibiting calcineurin is the rigidification of FK506. Consistent with this analysis are the reduced calcineurin inhibition of the H87V, R42K single mutants and the R42K+H87V double mutant of FKBP12. While weaker interactions between the FKBP12 mutant residues 42 and 87 with CaN may be responsible for the decreased calcineurin inhibition, it is likely that the reduced calcineurin inhibition is due, at least in part, to decreased conformational homogeneity of FK506 in its complex with the mutant FKBP.

F. FKBP12–Rapamycin–FRAP Complex

The binding of FRAP to the FKBP12−rapamycin complex (K_d = 2 nM) is mediated by a small (12-kD) domain of FRAP, the FKBP12−rapamycin binding (FRB) domain. The crystal structure of the ternary complex between FKBP12, rapamycin and FRB has been reported by a collaborative effort between Harvard and Cornell groups.[467]

In the ternary complex, rapamycin interacts extensively with both FKBP12 and FRB. The interactions between rapamycin and FRB are similar to those in the binary complex. There are no hydrogen bonds between rapamycin and FRB. The FRB domain contains a hydrophobic pocket which is formed by several aromatic residues. Rapamycin interacts with FRB through close VDW contacts with aromatic residues, and a series of interactions along the triene arm of rapamycin involving Phe-2039, Trp-2101, Tyr-2105, and Phe-2108 appear especially important. Ser-2035, Leu-2031, Thr-2098, Asp-2102, and Tyr-2038 also make contact with rapamycin. There is a small, but important, conformational change between the binary and ternary complexes. In the binary complex the triene arm is planar with the three double bonds fully conjugated. In the ternary complex bond rotations disrupt conjugation; this avoids a close contact with Phe-2108 and deeply buries the C23 methyl group of rapamycin in a small crease between Phe-2108 and Leu-2031.

Although rapamycin interacts extensively with both FKBP12 and FRB in the ternary complex, the extent to which the two proteins interact with each other is rather limited. Two regions of the complex show interactions between the proteins: the 40′s loop of FKBP12 with the α4 of FRB, and the 80′s loop of FKBP12 with the α1-α2 region of FRB. While the protein−protein interactions appear to be nonspecific, they do involve about 400 Å^2 of solvent-accessible surface area. A significant conformation change in FKBP12 occurs in the 80′s loop upon ternary complex formation. This conformational change, in the vicinity of Ile-90, moves FKBP12 away from FRB to avoid a steric repulsion.

The structure of the FKBP12, rapamycin, FRB ternary complex provides another example of how a small molecule ligand can mediate protein dimerization. This structure, like the FKBP12, FK506, calcineurin structure, shows extensive interactions between the ligand and each protein partner. In each case, hydrophobic contacts dominate over polar contacts. These two ternary complexes show less specific

protein−protein contacts; however, in each case the protein−protein contact do involve a significant amount of surface area.

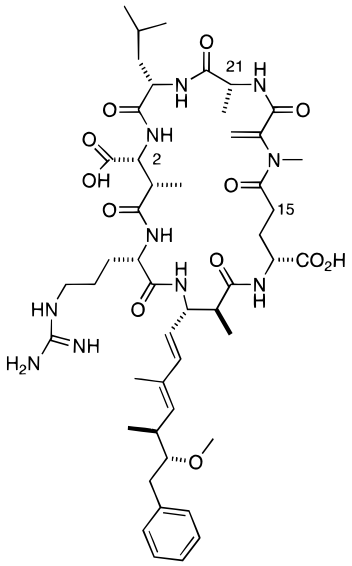
G. Direct Inhibitors of Serine/Threonine Phosphatases

The structural studies on the FKBP12−FK506−calcineurin complex also revealed, for the first time, the structure of a protein serine/threonine phosphatase. Another report has described the structure of a related phosphatase PP-1.[468] These structures have also introduced structural information about the active sites of these proteins and structures of inhibitors bound within the active site.

The phosphatase active sites for both CaN and PP-1 are quite homologous and are very similar in structure. The active site contains two metal atoms, a Zn^{2+} and an Fe^{3+}, in close proximity. These metal ions are ligated by histidines, aspartates, and an asparagine. The vacant coordination sites of the metals are filled by water molecules. A water molecule (or hydroxide ion) is observed which bridges the two metal ions. On the basis of the structure of the active site, several mechanistic proposals for phosphate hydrolysis have been put forth. Each of these proposals suggest a different water molecule as the nucleophile for phosphate hydrolysis. So far there is little direct experimental evidence for any particular reaction mechanism for these phosphatases. Resolution of correct mechanism will require additional structural studies with substrate based inhibitors.

A portion of the C-terminus of calcineurin contains residues which act as an autoinhibitory domain. For calcineurin, a structure of its active site inhibited by the calcineurin auto inhibitory (AI) domain (CaNA 469−486) has been described.[469] A 20-residue polypeptide, corresponding to CaNA residues 467−486, shows a $IC_{50} = 20\ \mu M$ as a competitive inhibitor of CaN.[470] The AI domain lies over the apparent substrate binding cleft in the catalytic domain. The most interesting interaction involves the carboxylate of Glu-481 with the catalytic groups in the active site. The carboxylate interacts with the Fe^{3+} in an indirect fashion. Rather than making direct VDW contacts with the metal, the carboxylate makes water mediated contacts with the Fe^{3+} ion. In addition, it makes a hydrogen bond with the side chain of Arg-254 of CaNA, the main chain carbonyl of Glu-481 also makes a hydrogen bond with the side chain of Arg-254. Surprisingly, there is only one hydrogen bond between the main chain of the AI-domain and the main chain of the catalytic domain of CaNA (CaNA-483 NH with CaNA 254 CO). The side chain of Asp-477 of the AI-domain makes hydrogen bonds with the phenol hydroxyl of Tyr-315 and the guanidine side chain of Arg-122. The only other hydrogen bond is between the main chain carbonyl of Pro-484 and the imidazole side chain of His-155. AI-domain residues Phe-470, Met-483, Pro-484, and Pro-485 make extensive hydrophobic contacts with CaN.

The active sites in the near vicinity of the metals are nearly identical for CaN and PP-1. However, the structure of these two proteins are different in regions adjacent to the metal sites. These differences



Microcystin-LR (**IM-34**)

Figure IM*26. Structure of natural product phosphatase inhibitor microcystin.

are presumably responsible for different substrate specificities. For PP-1, the natural product microcystin-LR (**IM-34**, Figure IM*26) is bound to the active site. Microcystin interacts with three distinct regions of the surface of PP-1. At the metal binding site, the C13 carboxylate interacts with the Fe^{3+} in an indirect fashion. Rather than making direct VDW contacts with the metal, the carboxylate makes water mediated contacts with the Fe^{3+} ion. This interaction is similar to the water-mediated carboxylate Fe^{3+} interaction seen in the CaN AI-domain. In addition, this carboxylate makes a hydrogen bond with the side chains of Tyr-272 and Arg-96. The C11 carbonyl makes a hydrogen bond with the side of Arg-221. A similar hydrogen bond is observed in the CaN AI structure. The C2 carboxylate makes hydrogen bonds to side chains of Tyr-134 and Arg-96. The $\alpha-\beta$ unsaturated amide near C18 has undergone a conjugate addition with the side chain of Cys-273, and the ligand is covalently attached to the protein. This covalent interaction is not essential for microcystin inhibition. The long hydrophobic side chain at C9 packs into a hydrophobic groove and makes intimate VDW contacts. The conformation of microcystin in solution is quite similar to its structure in the complex with PP-1. Thus, the ligand is preorganized for binding; this contributes toward its high affinity.

H. Summary and Perspective

Studies in the immunophilin area have demonstrated, for the first time, that small molecule ligands are capable of dimerizing two proteins which have no affinity for each other. Structural work in this field has demonstrated, on the atomic level, how this phenomenon occurs. The structures of the ternary complexes of the FKBP12−FK506−calcineurin complex and the FKBP12−rapamycin−FRB structures have shown that the major forces holding each of these complexes together are hydrophobic contacts of the ligands with each protein partner. The structure of the cyclophilin−cyclosporin−calcineurin ter-

nary complex is now needed to make the immunophilin story complete.

The binary complexes of FK506−FKBP12 and rapamycin−FKBP12 both show that the ligands expose a hydrophobic surface area into solvent with a well-defined geometry. It is this exposed hydrophobic surface which is observed to make hydrophobic contacts with the third partner of the ternary complex and is the "molecular glue" responsible for making each ligand a "chemical inducer of dimerization".

In addition, this section has described several binary complexes with FKBP12 which are not chemical inducers of dimerization. These defective protein−ligand complexes are not capable of exposing hydrophobic surface area into solvent with a well defined geometry. The structures of these binary complexes provide insight into why these complexes are not capable of "double-edged" molecular recognition properties. In many cases, unproductive hydrophobic collapse is responsible, while in others it is an increased conformational heterogeneity of the ligand in the binary complex.

Dimerization and oligomerization are general biological control mechanisms contributing to the activation of cell membrane receptors, transcription factors, vesicle fusion proteins, and other classes of intra- and extracellular proteins.[422] Recent studies have shown that artificial systems can be designed to undergo dimerization upon addition of an appropriate ligand which can act as a chemical inducer of dimerization for that system. This method has potential to be applied to signal transduction activation and termination pathways and has long term potential in gene therapy.[471] The structural studies described in this section, as well as others yet to be achieved, have potential in assisting the growth of this new area of research.

VIII. Other Drug Design Targets: Brief Summary

A. Inhibitors of Folate-Binding Enzymes

Folic acid (**XL-1**) is an essential vitamin that, in its reduced form, acts as a cofactor to many metabolic enzymes by carrying and facilitating the transfer of one-carbon units.[472] Many folate-dependent enzymes participate in metabolic pathways that are critical to the life cycle of cells. As a result inhibitors of these classes of enzymes have potential as anticancer agents.[473,474] Historically, folate-binding enzymes have played an important role in the development of structure-based drug design strategies.

Methotrexate (**XL-2**, Figure XL*1), discovered in 1948 at Lederle,[475] is an inhibitor of the folate-dependent enzyme dihydrofolate reductase (DHFR) and is widely used for cancer chemotherapy in humans. DHFR is a small, stable, and crystallographically well-behaved enzyme. As a result, this enzyme was the subject of many of the early structural studies of protein−ligand complexes. These pioneering retrospective analyses provided the foundation[476,477] for the development of the iterative protein structure-based design process.[12] These structural studies of DHFR with various ligands has been

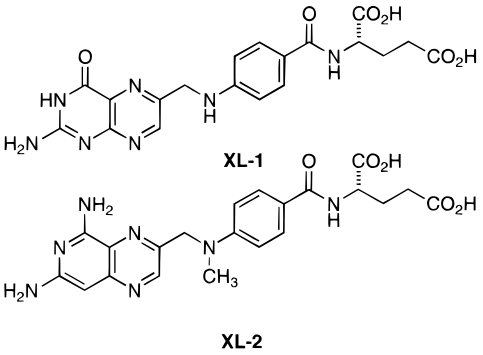


Figure XL*1. Structures of various compounds.

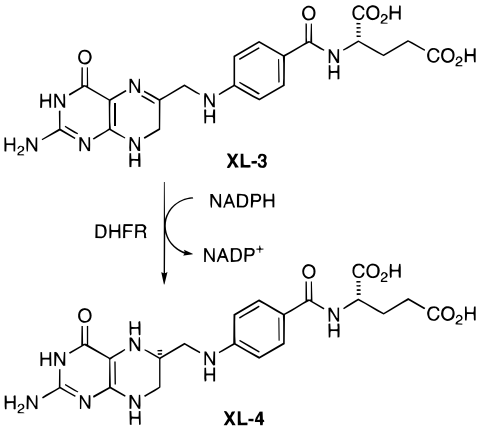


Figure XL*2. DHFR catalyzes the reduction of dihydrofolate to tetrahydrofolate.

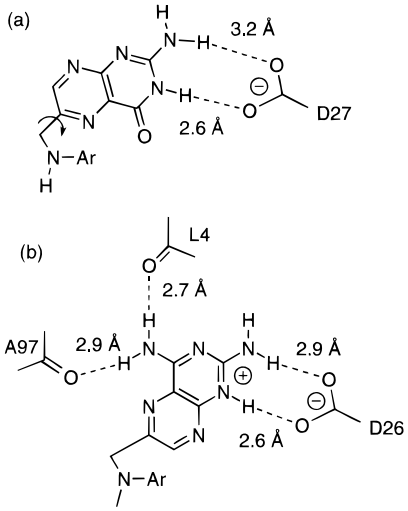


Figure XL*3. (a) Folic acid interacts with DHFR in a substrate like fashion; however, by a single bond rotation methotrexate (b) binds with a reversed orientation of its pteridine ring.

previously reviewed; [478−481] therefore, only a few important aspects of that work will be discussed.

DHFR catalyzes the NADPH-dependent reduction of the 5,6 double bond of dihydrofolate (**XL-3**) to produce tetrahydrofolate (**XL-4**). (See Figure XL*2.) This activity is responsible for maintaining the folate cofactor in its fully reduced and biologically relevant oxidation state. The ternary complex between DHFR, NADPH, and methotrexate has been reported (PDB entry: 3DFR),[482] as well as the ternary complex between DHFR, $NADP^+$, and folic acid (PDB entry: 7DFR).[483] While **XL-1** and **XL-2** are close structural

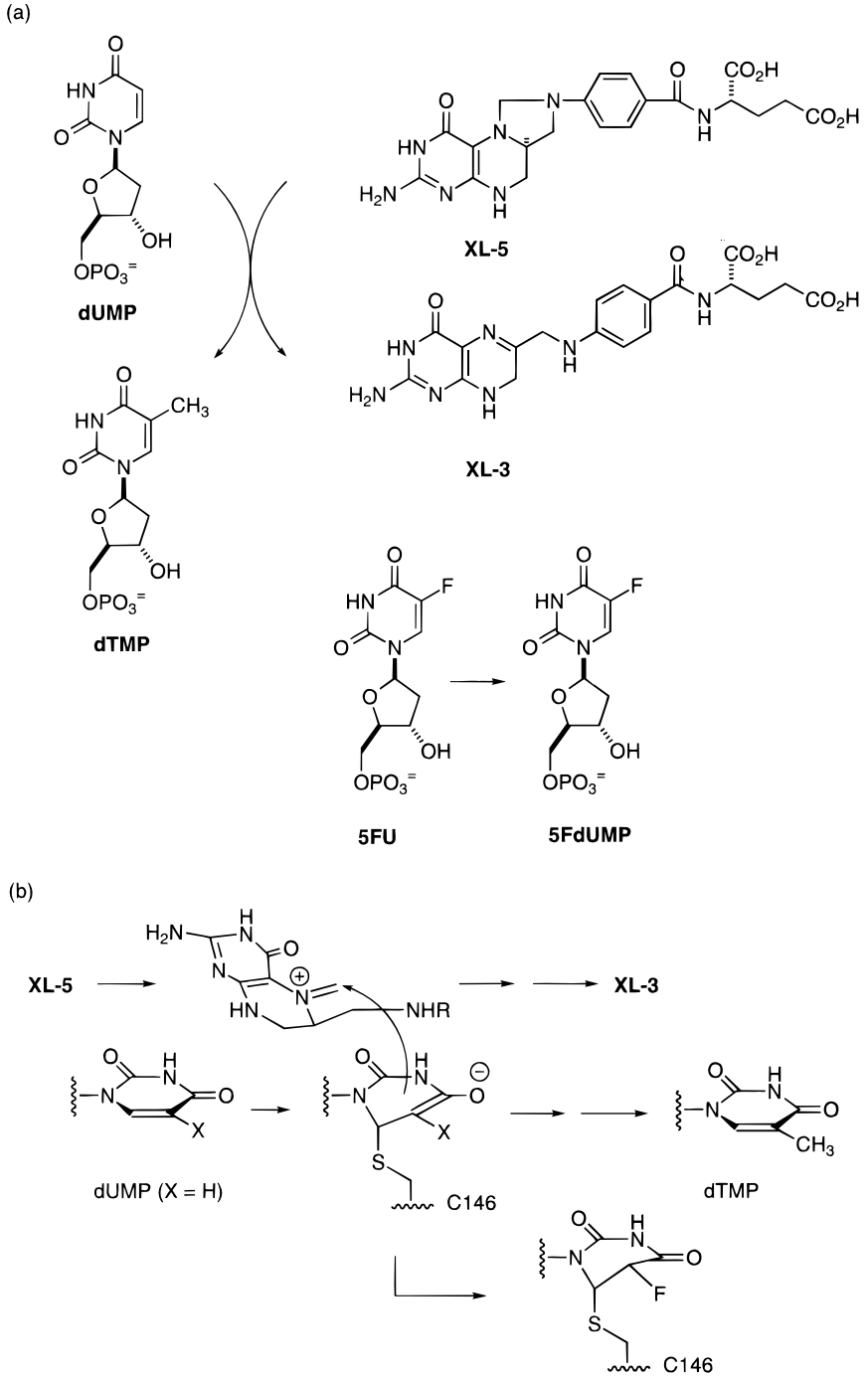


Figure XL*4. (a) Thymidylate synthase (TS) catalyzes the conversion of deoxyuridate (**dUMP**) to thymidylate (**dTMP**) using 5,10-methylene tetrahydrofolate (**XL-5**) as a cofactor. Compound **5FU** is a prodrug for the mechanism-based TS inhibitor **5FdUMP**. (b) The mechanism of the one-carbon atom transfer catalyzed by TS involves thiol conjugate addition to generate an enolate, followed by enolate addition to an open form of the folate cofactor. The fluorine atom of **5FdUMP** prevents proper processing, a covalent enzyme inhibitor complex results.

analogs, they bind quite differently to DHFR. The two ligands differ in by an approximately 180° degree bond rotation which results in different orientations of the pteridine rings of the two ligands. Compound **XL-1** binds in a substrate like fashion, making hydrogen bonds to Asp-27. The higher affinity ligand **XL-2** binds with a different orientation of its pteridine ring and makes more extensive hydrogen bonds with the protein. The protonated pteridine ring makes hydrogen bonds with both Asp-26 and the backbone carbonyls of Leu-4 and Ala-97. These interactions are only possible with an orientation of this ring which is different than that of a substrate-like analog. (See Figure XL*3.) The different binding orientations of **XL-1** and **XL-2** was an early illustration of the importance of obtaining cocrystal structures of different ligands bound to the same protein as a prelude to further structure based design.

Thymidylate synthase (TS) catalyzes the conversion of deoxyuridate (**dUMP**) to thymidylate (**dTMP**) using 5,10-methylene tetrahydrofolate as a cofactor. 5-Fluorouracil (**5FU**), which is a prodrug for **5FdUMP**, is a potent mechanism-based inhibitor of TS which is used clinically in the treatment of cancers. The catalytic mechanism of TS involves

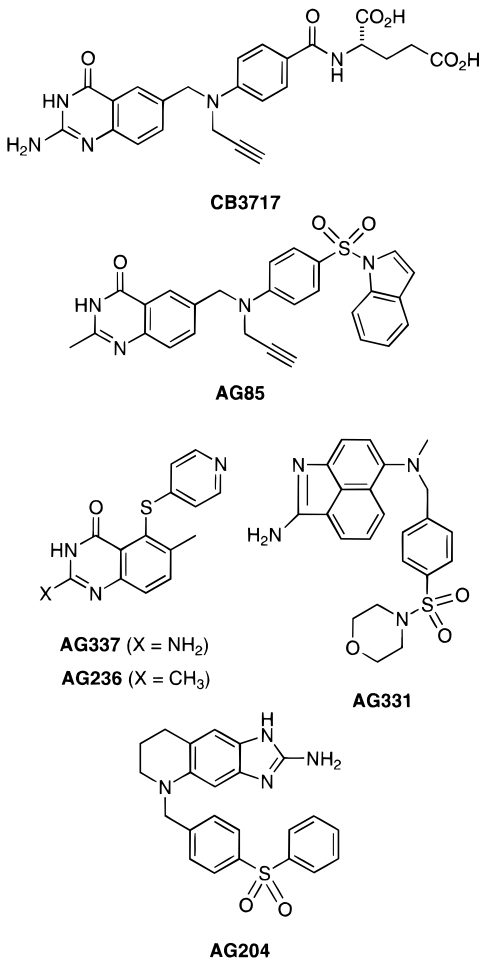


Figure XL*5. Structures of various TS inhibitors.

conjugate addition of a thiol (Cys-146) to the enzyme bound **dUMP**, to give an enolate. The enolate is then alkylated by the iminium ion form of the folate cofactor **XL-5**. Further chemistry results in the products **dTMP** and folate **XL-3**. The presence of the fluorine atom in **5FdUMP** prevents appropriate proton transfers and results in the irreversible addition of the side chain of Cys-146 to the inhibitor. (See Figure XL*4.) A detailed discussion of the catalytic mechanism of TS and its inhibition by **5FdUMP**, based on the structure of the ternary inhibited complex, has been published.[484]

In 1991, a seminal perspective article[12] was published by the Agouron group describing the design of thymidylate synthase (TS) inhibitors using iterative protein crystallographic analysis in a prospective manner. This work, which has been discussed in short review articles,[13,485] has resulted in four structurally diverse, potent TS inhibitors (**AG85**, **AG337**, **AG331**, and **AG204**). Three of these compounds (**AG85**, **AG337**, and **AG331**) have advanced into human clinical trials. Each inhibitor binds to the folate binding site of TS. A brief discussion of how each of these structurally diverse inhibitors bind to the same site of the enzyme follows. (See Figure XL*6.)

Compound **AG85** (K_i = 13 nM) is an analog of a known, potent TS inhibitor **CB3717** (K_i = 4.8 nM).[486] The design objective was not to generate a more potent enzyme inhibitor, but rather to alter the physical and pharmacological properties of **CB3717** by replacing the glutamate side chain. The structure of the ternary complex between *E. coli* TS, **5FdUMP**, and **CB3717** (PDB entry: 2TSC)[487] was used as the starting point in the design process. The quinazolinone ring of **CB3717** binds at the base of a large funnel-shaped cavity. One face of the quinazolinone ring makes VDW contacts with the uridine and sugar ring of **5FdUMP**, the edge of the ring and the propargyl group make VDW contacts with the protein, especially the side chains of Ile-79, Trp-80, Trp-83, Ala-263, Val-262, and Tyr-209. The aniline ring fits into a narrow cleft and makes VDW contacts with the side chains of Ile-79, Leu-172, and Phe-176. There is an overall high degree of steric complementarity between the ligand and the TS−**5FdUMP** complex. The glutamate side chain of **CB3717** is somewhat solvent exposed. The quinazolinone ring makes a hydrogen bond to the side chain of Asp-169 and a hydrogen bond to the backbone carbonyl of Ala-263. In addition, there is a water molecule which makes hydrogen bonds to the backbone carbonyl of Ala-263, the side chain of Arg-21 and the ring nitrogen of **CB3717**. The ring system of **AG85** differs from **CB3717** in the substitution of an NH_2 group with a methyl group. This results in a loss of a hydrogen bond to the carbonyl of Ala-263. The ring system of **AG85** binds nearly identically to **CB3717**. The indole ring of **AG85** is disordered and the sulfonyl group makes VDW contacts with the side chain of Leu-172. The sulfonyl group does not make any direct close electrostatic interactions with the protein. The altered physical and pharmacological properties of **AG85** warranted its evaluation in human clinical trials.

The structure of **AG236** (K_i = 500 nM), a close analog of **AG337** (K_i = 15 nM), bound to TS,[488] shows that the quinazolinone ring binds in nearly the same fashion as **CB3717** and **AG85**. The 6-methyl group improves steric complementarity and contributes significantly to its potent activity. The pyridyl ring fits into the cavity that the benzene rings of **CB3717** and **AG85** bind to. Substitution of the 2-methyl group with an amino group, to give **AG337**, improves potency by a factor of 33; this is presumably due to formation of hydrogen bonds to Ala-263 and the structural water molecule. Compound **AG337** is now in advanced clinical trials as an anticancer agent.

Compound **AG331** (K_i = 12 nM) is a structurally novel TS inhibitor which was designed *de novo*.[489] The nathylstyril ring system binds at the base of the active site cleft and one face of the ring system makes VDW contacts with the uridine and sugar ring of **5FdUMP**, the edge of the ring makes VDW contacts with the protein, especially the side chains of Ile-79, Trp-80, Trp-83, Ala-263, Val-262, and Tyr-209. Thus, using a very different ring system, **AG331** fills the same space as **CB3717** and **AG236** do. The amidine group of **AG331**, which is presumably protonated, makes hydrogen bonds to the side chain of Asp-169 and to a structural water molecule. The phenyl ring fits into the cavity that the benzene rings of **CB3717** and **AG85** bind to. The sulfonyl group and the morpholine group are solvent exposed and make some VDW contacts with the protein. The potent activity and desirable properties of **AG331** warranted

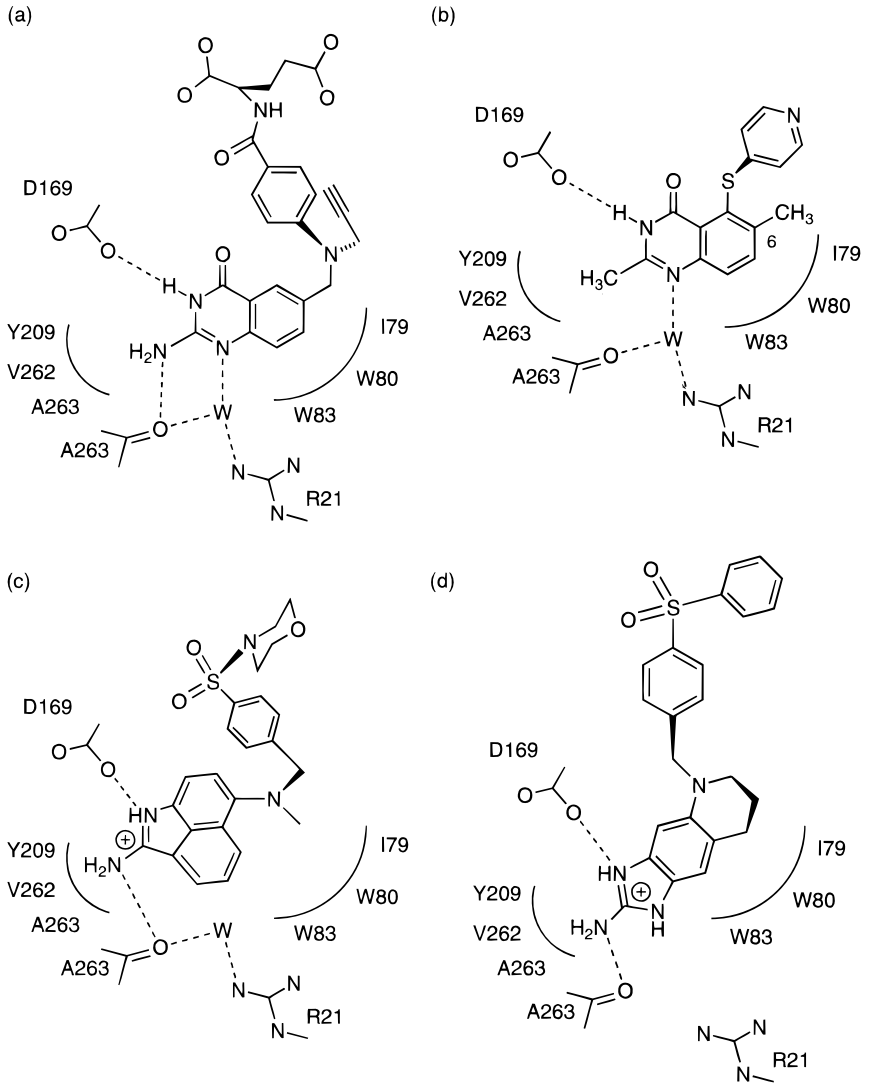


Figure XL*6. Interactions between TS and inhibitors (a) **CB3717**, (b) **AG85**, (c) **AG236**, (d) **AG331**, and (e) **AG204**.

its evaluation in clinical trials as an anticancer agent.

Compound **AG204** (K_i = 500 nM) is another structurally novel TS inhibitor which was designed *de novo*.[490] The imidazotetrahydroquinoline ring system binds at the base of the active-site cleft and one face of the ring system makes VDW contacts with the uridine and sugar ring of **5FdUMP**, the edge of the ring makes VDW contacts with the protein, especially the side chains of Ile-79, Trp-80, Trp-83, Ala-263, Val-262, and Tyr-209. Thus, **AG204**, **AG337**, and **AG331** fill approximately the same active site space but do so with quite different ring systems. The protonated guanidine group of **AG204**, like **SB3717** and **AG331**, forms hydrogen bonds to Asp-169 and Ala-263. Unlike **SB3717** and **AG331**, **AG204** does not interact with the structural water molecule, and no crystallographic evidence for this water molecule was observed. The failure to trap this water molecule is a favorable entropic process and may contribute to inhibitor potency. Analogs of **AG204**, designed using the iterative cycle, were significantly more potent than **AG204** as TS inhibitors.[490]

These examples illustrate how three structurally distinct ring systems are capable of binding to the same active site. Each ring system makes similar electrostatic interactions with the protein, fill approximately the same space, and make similar VDW contacts. (See Figure XL*6.)

Another important folate-dependent enzyme is phosphoribosylglycinamide formyltransferase (GART). GART catalyzes the transfer of a formyl group from 10-formyltetrahydroformate (**10fTHF**) to glycinamide ribonucleotide (GAR), a reaction in the purine biosynthetic pathway. An antifolate, 5,10-dideazatetrahydrofolate (**5fTHF**),[491,492] has been shown to inhibit GART, *de novo* purine biosynthesis, and proliferation of tumor cells in culture[493] and has been evaluated in human clinical trials. The structure of the ternary complex of GART, GAR, and **5fTHF** has been reported (PDB entry: 1CDE).[494] (See Figure XL*7.)

The bicyclic ring system of the inhibitor **5fTHF** binds to a hydrophobic region of the protein and makes several hydrogen bonds to backbone groups of the protein. The aminomethyl substituent is axially disposed on the bicyclic ring system and the phenyl ring makes close hydrophobic contacts, the glutamate side chain is mostly solvent exposed.

B. Carbonic Anhydrase Inhibitors

Human carbonic anhydrase (HCA) catalyzes the diffusion-controlled hydration of carbon dioxide to

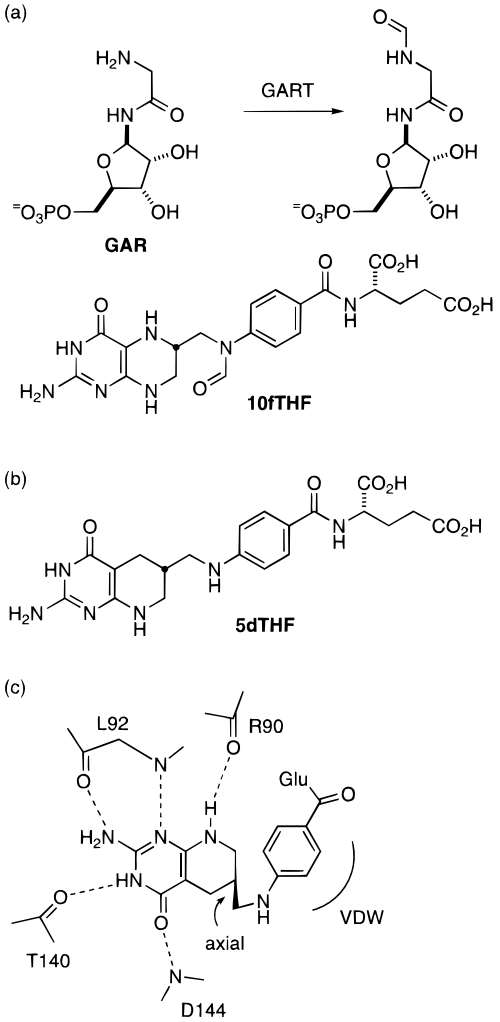


Figure XL*7. (a) GART catalyzes the transfer of a formyl group from the cofactor **10fTHF** to the substrate **GAR**. (b) Structure of GART inhibitor **5dTHF**. (c) Interactions between **5dTHF** and GART.

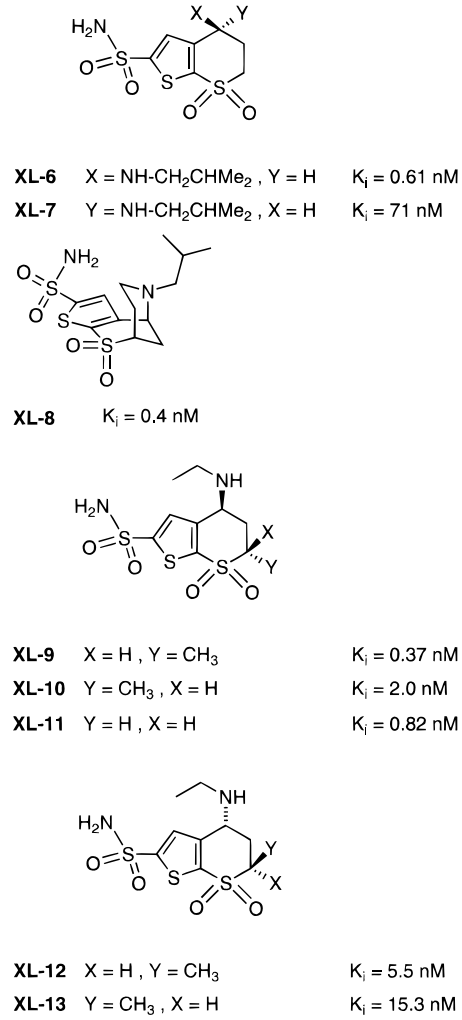


Figure XL*8. Structures of carbonic anhydrase (HCA) inhibitors.

bicarbonate ($k_{cat}/K_m = 1.5 \times 10^8$ M^{-1} s^{-1}). Inhibitors of HCA can influence fluid dynamics in the eye and be useful in the treatment of glaucoma.[495] The Merck group has reported structure-based design of inhibitors using the structure of HCA.[13,496–498] The active site contains a zinc atom which is tetrahedrally coordinated by three histidines. The active-site cavity is amphiphilic; one wall is composed of hydrophilic residues and the other by hydrophobic residues.

The enantiomeric compounds **XL-6** and **XL-7** were both found to be potent inhibitors of HCA, **XL-6** being more than 100-fold more potent than **XL-7**. Structures of the complexes between HCA and **XL-6** and **XL-7** have been determined by X-ray crystallography. For both enantiomers, the sulfonamide group is coordinated to the active-site zinc atom through the presumably deprotonated sulfonamide nitrogen. For both enantiomers, the isobutyl amine side chain is axially disposed on the six-membered ring and makes VDW contacts with the same region of the protein. To accomplish this, different conformations of the linking group, including a 20° rotation around the NSCS torsion angle and a rotation from anti to gauche of the isobutyl amino side chain, are required. Thus, a major contributor to the difference in potencies of the two inhibitors has been attributed to higher ligand strain in the less potent **XL-7**.

To overcome the conformational preference of the ring and lock the amino side chain in an axial conformation, the constrained tricyclic analog **XL-8** (Figure XL*8) was prepared. This compound was found to be nearly twice as potent as an HCA inhibitor.[496]

A second approach to overcoming the conformational preference of **XL-7** has been reported.[13] Since the more potent enantiomer **XL-7** binds to the protein with an axial substituent, the six-membered ring was substituted with a methyl group. It was expected that **XL-9**, with a methyl group trans to the ethylamine, the energetic penalty for having the aminoethyl group axial would be decreased. For **XL-9**, either the methyl group or the aminoethyl group must be axial; therefore, there is not a significant ligand conformational strain penalty for having an axial aminoethyl group in the bound ligand. (See Figure XL*9a.) It was observed that **XL-9** showed a 2.2-fold increase in potency over **XL-11** (0.47 kcal mol^{-1}). The exchange of the isobutyl for ethyl group was to alter the physical properties of the molecule. Compound **XL-9** (dorzolamide) is now a marketed antiglaucoma agent.

The structure of the **XL-9** complex with HCA (PDB entry: 1CIL) reveals the following interactions. (See

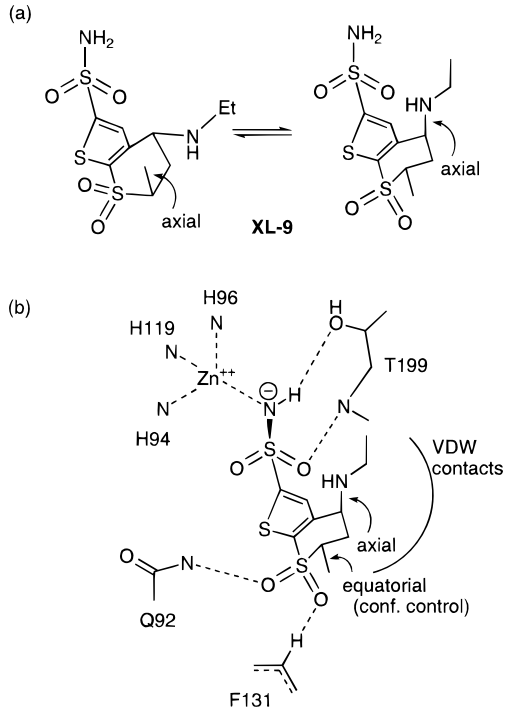
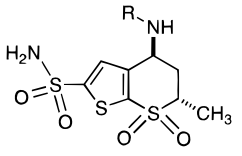


Figure XL*9. (a) The methyl group in **XL-9** serves as a conformational control element by making the axial aminoethyl substituent less unfavorable. (b) Key interactions of **XL-9** with HCA. The methyl substituent on the six-membered serves as a conformational control element, allowing the amino ethyl substituent to be axial and not suffer a major energetic penalty.

Figure XL*9.) The sulfonamide nitrogen, which is presumably deprotonated, ligates the zinc ion (1.95 Å) and displaces a water molecule observed in the apoenzyme (PDB entry: 1CA2),[499] it also makes a hydrogen bond to the side chain of Thr-199 (2.81 Å). The zinc is also ligated by imidazole nitrogens of three histidines of HCA (His-94, His-96, and His-119). One of the sulfonamide oxygens forms a hydrogen bond to the backbone NH of Thr-199 (3.0 Å). The NSCS dihedral angle is 144°. The bicyclic ring system of **XL-9** makes hydrophobic contacts with HCA, including the side chains of Val-121 and Leu-198. One of the sulfone oxygens makes a polar interaction with the side chain of Gln-92 (3.15 Å) and the other with an aromatic CH of Phe-131 (3.4 Å). The aminoethyl substituent is axial on the six-membered ring and the ethyl side chain makes VDW contacts with the protein, especially the side chains of Trp-5 and His-64. His-64, a key residue in the catalytic mechanism of HCA, has its side chain rotated 3.1 Å away from its position in the apoenzyme. The methyl group is equatorial on the six-membered ring does not make significant VDW contacts with the protein; however, it serves as a conformational control element by making the axial aminoethyl substituent less unfavorable.

As anticipated, the trans isomer **XL-9** was more potent than the cis isomer **XL-10**. The enantiomeric compounds **XL-12** and **XL-13** were also prepared and evaluated as HCA inhibitors and were, as expected, less potent. Crystal structures of HCA with all four isomers were determined.[13] A major difference in the conformation of the four isomers was the NCSC dihedral angle. It was proposed that a major contributor to the differences in affinity for each isomer was ligand torsion strain of this key NSCS bond.



XL-9	R = CH_2CH_3	K_i = 0.37 nM
XL-14	R = CH_3	K_i = 1.88 nM
XL-15	R = H	K_i = 1.52 nM

Figure XL*10. Structures of HCA inhibitors; the structures of their complexes show that the more potent **XL-9** moves the side chain of His-46 and displaces an enzyme-bound water molecule, **XL-14** and **XL-15** do not.

The three inhibitors **XL-9**, **XL-14**, and **XL-15** are all potent HCA ligands which differ in affinity by affinity by less than 1 kcal mol^{-1}. Structurally they differ in the amino substituent, the ethyl substituent being most potent and the methyl and hydrogen substituents being less potent and nearly equipotent to each other. (See Figure XL*10.) Structures for the HCA complexes of these three ligands have been reported (PDB entries: 1CIL, 1CIM, 1CIN).[498] For all three complexes the inhibitors bind nearly identically. The major difference observed is the conformation of the side chain of His-64 of HCA. For the less potent inhibitors **XL-14** and **XL-15** the position of this side chain is identical to that of the apoenzyme, while for the more potent **XL-9** it has is rotated away by 3.1 Å. The larger ethyl side chain of **XL-9** forces the histidine side chain to move away from its native position to avoid VDW repulsion. This side-chain movement results in a structural water molecule, observed in the apoenzyme and the complexes of **XL-14** and **XL-15** to be displaced into bulk solvent. It has been proposed that the increase in entropy when this water molecule is released contributes to the increase in binding and overcomes the small penalty for putting the His-64 side chain in a higher energy state.[498] Thus, the increased binding affinity of **XL-9**, relative to **XL-14** and **XL-15** appears to be due to entropic factors.

These studies are very interesting. They illustrate the importance of ligand conformation strain in the overall stability of the complex. The design of **XL-9**, which is more potent than **XL-11**, provides a nice example of introducing a methyl group to alter the conformational energetics of the bound isomer to improve binding affinity.

C. Inhibitors of Phospholipase A_2

Phospholipase A_2 (PLA2) catalyzes the *sn*-2 acyl-hydrolysis of phospholipids liberating free fatty acids, including arachidonic acid, and lysophospholipids. Since arachidonic acid is metabolized to form a variety of inflammatory agents inhibitors of PLA2 have potential as antiinflammatory agents.[500] The mechanism of ester hydrolysis by PLA2 involves nucleophilic attack by a water molecule which is activated by an aspartate−histidine dyad and a calcium ion (see Figure XL*11a).[501] The structure of a phosphate-based transition analog has bee reported (PDB entry: 1POE; see Figure XL*11b).[502]

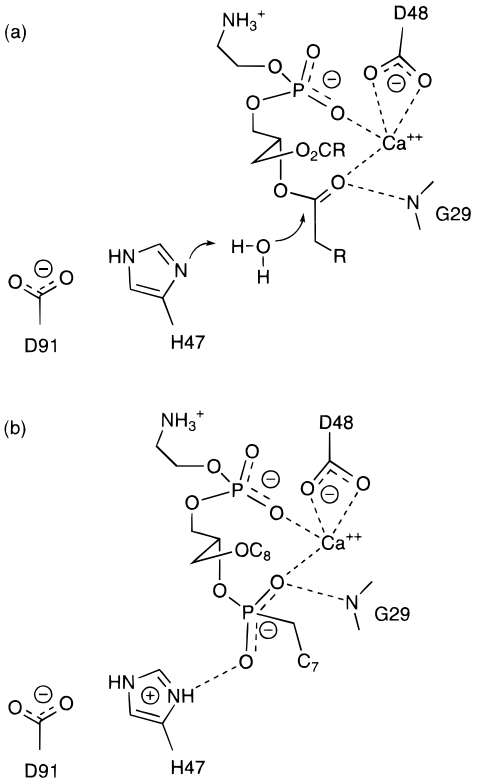
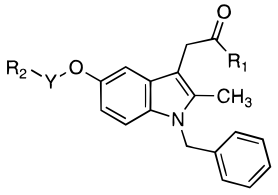


Figure XL*11. (a) Likely mechanism for ester hydrolysis by PLA2 involves nucleophilic attack by an enzyme-activated water molecule on the scissile ester bond with a Ca^{+2} ion acting as a Lewis acid and (b) structure of PLA2 with a transition state analog. The hydrophobic *n*-alkyl chains of the inhibitor bind in a deep hydrophobic channel.



	R_1	R_2	Y	$X_i(50)$
XL-16	OH	H	CH_2	.014
XL-17	NH_2	H	CH_2	.00068
XL-18	NH_2	CO_2H	$(CH_2)_3$	.00012
XL-19	NH_2	PO_3H_2	$(CH_2)_3$	.000045

Figure XL*12. Structures and affinities of various PLA2 inhibitors.

Indole **XL-16** (Figure XL*12) was identified as a PLA2 inhibitor during a random screening process. It had an $X_i(50)$ of 0.014 mole fraction (concentration in mole fraction required for 50% enzyme inhibition). The structure of the complex between **XL-16** and PLA2 was determined by X-ray crystallography. The indole ring of **XL-16** binds in the hydrophobic channel of PLA2. The enzyme undergoes a significant conformational change of side-chain residues to accommodate the inhibitor. The active-site calcium ion has been displaced in the complex and the R_1 carboxylate of the inhibitor interacts with Asp-49 which normally ligates the calcium ion. Clearly, either the inhibitor carboxylate or the Asp-49 carboxylate is protonated in the complex. The active site His-48 has also undergone a conformational change in the complex.

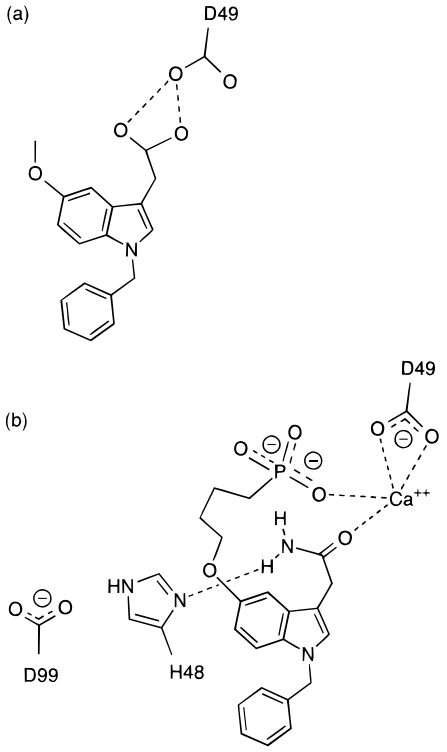
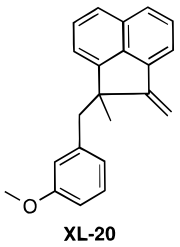


Figure XL*13. (a) Interaction of screening lead **XL-16** with PLA2. The hydrophobic core of the inhibitor binds to a hydrophobic channel of the enzyme, large side chain movements by the enzyme are needed to accommodate the inhibitor. The carboxylate of the inhibitor makes a bifurcated hydrogen bond to Asp-49, displacing the calcium ion. There is no interaction with His-48. (b) Interaction of potent inhibitor **XL-19** with PLA2. Substitution of the carboxylate with an amide allows for ligation of the calcium ion and hydrogen bond formation with His-48. The phosphonate side chain allows for additional ligation of the calcium ion.

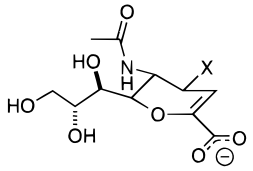
On the basis of the structure of the **XL-16** complex it was proposed that conversion of the R_1 carboxylate to a primary amide would allow the calcium ion to be ligated and allow the inhibitor to form a hydrogen bond with His-48. Inhibitor **XL-17** was prepared and found to be 20 times more potent than **XL-16** as a PLA2 inhibitor. The structure of the complex between **XL-17** and PLA2 showed that the design predictions were correct. The primary amide group of the inhibitor and the side chain of Asp-49 now ligated a calcium ion and the inhibitor formed a hydrogen bond to His-48. Elaboration of the methoxy group of the inhibitor resulted in **XL-18** and **XL-19**, crystal structures of their complexes showed that the carboxylate and phosphonate groups, respectively, acted as calcium ion ligands. The structure-based modifications of screening lead **XL-16** resulted in phosphonate **XL-19** which is 6,400 times more than the original screening lead. (See Figure XL*13.) This work[503] is a nice example of protein structure-based inhibitor optimization.

In 1987 the duPont group reported the design, synthesis and evaluation of PLA2 inhibitor **XL-20** (Figure XL*14) using the crystal structure of bovine pancreatic PLA2.[504] This compound has an IC_{50} of 0.5 μM as an inhibitor of PLA2. This was one of first, dramatic reports of *de novo* design based purely upon protein structure. While SAR data of analogs is consistent with the proposed binding of **XL-20** to the



XL-20

Figure XL*14. Structure of PLA2 inhibitor **XL-20**.



XL-21	X = OH	K_i = 1000. nM
XL-22	X = NH_2	K_i = 50. nM
XL-23	X = NHC(NH_2)=NH	K_i = 0.2 nM

Figure XL*15. Structures of sialidase inhibitors.

hydrophobic channel of PLA2, no reports have described the structure of the complex.

D. Sialidase Inhibitors

Crystal structures of two influenza virus surface proteins, hemagglutinin[505,506] and sialidase,[507,508] have been reported and are potential targets for structure-based drug design. The influenza virus sialidase is an enzyme which cleaves terminally α-ketosidically linked sialic acids from glycoproteins, glycolipids, and oligosaccharides. It has been proposed that this enzyme helps the elution of newly synthesized virons from infected cells.[509]

The structure of the complex between influenza virus sialidase (IVS) and the 1 μM inhibitor **XL-21** has been reported (PDB entry: 1NSD; Figure XL*15).[510] The enzyme active site is very charged and hydrophilic having five arginine residues, five glutamic acid residues, and one aspartic acid residue in the active-site region. The syn carboxylate oxygens of **XL-21** forms a strong salt bridge with Arg-373 and the anti oxygens form weaker salt bridges to Arg-115 and Arg-291. The triol side chain interacts with Glu-274 and Glu-275, the acetamide oxygen interacts with Arg-149 and the acetamide NH interacts through a structural water molecule with Glu-225 and Glu-275. The 4-hydroxyl group has a water-mediated interaction with Glu-225 and Glu-116. (See Figure XL*16.)

The structure of the complex between **XL-21** and IVS was the starting point for a short, but highly efficient, drug design effort.[511] Analysis of the active site, with the assistance of the software program GRID,[512] suggested replacing the 4-hydroxyl group of **XL-21** with a protonated primary amine group. This was expected to result in salt bridge formation with the side chain of Glu-116 (PDB entry 1NSD numbering). Compound **XL-22** was prepared and found to be 20 times more potent as a sialidase inhibitor than **XL-21**. The structure of the complex between **XL-22** and IVS, as expected, showed a salt bridge between the 4-amino group and the carboxy-

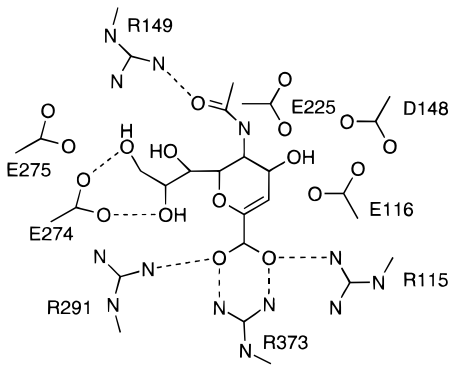


Figure XL*16. Interactions of **XL-21** with the charged and polar active site of IVS.

late of Glu-116. It was further suggested that replacement of the amino group of **XL-22** with a more basic guanidino group would result in a further increase in potency. The guanidine analog **XL-23** was prepared and found to be 250 times more potent than the amine **XL-22**. The structure of the complex between **XL-23** and IVS showed that the newly introduced guanidine group interacts with both Glu-116 and Glu-225. Since IVS is a cell surface protein, the zwitterionic **XL-23** does not need to cross a cell membrane to reach its target. Compound **XL-23** showed activity as an antiinfluenza agent.

E. Purine Nucleoside Phosphorylase Inhibitors

Structure-based design of purine nucleoside phosphorylase inhibitors has been reviewed recently.[513]

F. Combinatorial Synthesis and Structure-Based Combinatorial Chemistry

Many tyrosine kinase signal transduction pathways rely on two small protein modules known as SH2 (Src homology 2) and SH3 (Src homology 3) domains. The key aspect of the function of SH2 and SH3 domains is their ability to recognize particular amino acid sequences in their target proteins. SH2 domains bind tightly to phosphorylated tyrosine residues. Structures of complexes between SH2 domains and phosphorylated peptides have been determined and these interactions have been previously reviewed.[514] In contrast, SH3 domains bind to nonphosphorylated peptide sequences that are rich in proline and hydrophobic residues. The lack of natural high-affinity ligands to SH3 domains delayed the determination of the structure of an SH3 ligand complex.

The screening of collections of compounds (libraries) against protein targets is a well-known method for the discovery of new leads. Many of the examples in this review have shown how natural product leads, and leads from random screening of large collections of pure compounds have been starting points for structure based design programs. The recent development of technologies for the preparation and screening of large combinatorial libraries[515] is likely to be an important new source for new leads in structure based design programs.

One SH3 domain is present in phospatidylinositol 3-kinase (P13K). The specific role of the P13K SH3 domain in signal transduction pathways is not well

understood. No natural ligands for these proteins are known. Recent studies by the Harvard and Ariad groups have reported the discovery of P13K SH3 ligands, by screening combinatorial libraries, and the determination of the structure of their protein–ligand complexes.[516]

Two different random peptide libraries, one containing two million hexapeptides and another containing 2 million cyclic heptapeptides, each covalently attached to a unique bead, [517] were prepared. These two large libraries, on beads, were screened for binding to the P13K SH3 protein. Both screenings failed to identify any SH3 ligands. On the basis of subsequent reports that SH3 domains appear to bind to proline rich motifs (PPXP), a biased combinatorial peptide library, XXXPPXPXX, where X was any amino acid except cysteine, was designed and prepared. This nonapeptide library of approximately 2 million members, covalently attached to beads, was then screened for its ability to bind to the P13K SH3 protein. Seventeen of the beads with the highest affinity for the protein were then isolated and sequenced by Edman degradation. This biased combinatorial library resulted in the discovery of two classes of ligands. One class (class I, 13 members) had the consensus sequence RXLPPRPXX. The second class (class II, 4 members) had the consensus sequence XXXPPLPXR. Six representative peptides were chemically synthesized and their affinities for the native SH3 domain were measured. The observed K_d's ranged from 8.7 to 30 μM. A representative class I peptide, RKLPPRPSK, had $K_d = 8.7$ μM, and a class II peptide LNKPPLPKR, had $K_d = 13$ μM. These were among the highest affinities of ligands for SH3 domains reported at that time.

Screening these libraries against the related Src SH3 protein found a similar trend. Two distinct classes of ligands were discovered. The major difference was that for the Src SH3 protein the class I ligands had the consensus sequence RXLPPLPXX, where a leucine replaced an arginine at position 6. A mutational analysis for these classes of peptide ligands provided addition SAR information. These studies provided ligands for structure determination, by NMR, of a representative member of each of these ligands in complex with the Src SH3 domain protein.

The NMR structures of the complex between the class I peptide ligand RALPPLPPY ($K_d = 8.0$ μM; PDB entry: 1RLQ) and the class II peptide ligand AFAPPLPRR ($K_d = 59.0$ μM; PDB entry: 1PRM) and the Src SH3 domain protein has been reported by the Harvard group.[518] Both ligands adopt a left-handed polyproline type II helix and make a salt bridge between an arginine side chain and the conserved Asp-99 of the SH3 domain. The other protein–ligand contacts in both complexes are mostly nonspecific VDW contacts between hydrophobic groups on side chains. There are no hydrogen bonds between the inhibitor backbone and the protein backbone. The difference between class I and class II peptide ligands is that they bind to SH3 domains in reverse orientations, that is, the amino to carboxyl directionalities of their helices are opposite. (See Figure XL*17.) The peptide orientation is determined by a salt bridge which is formed with Asp-99. In addition the structural studies showed that the C-terminus of class I ligands and the N-terminus of class II are solvent exposed.

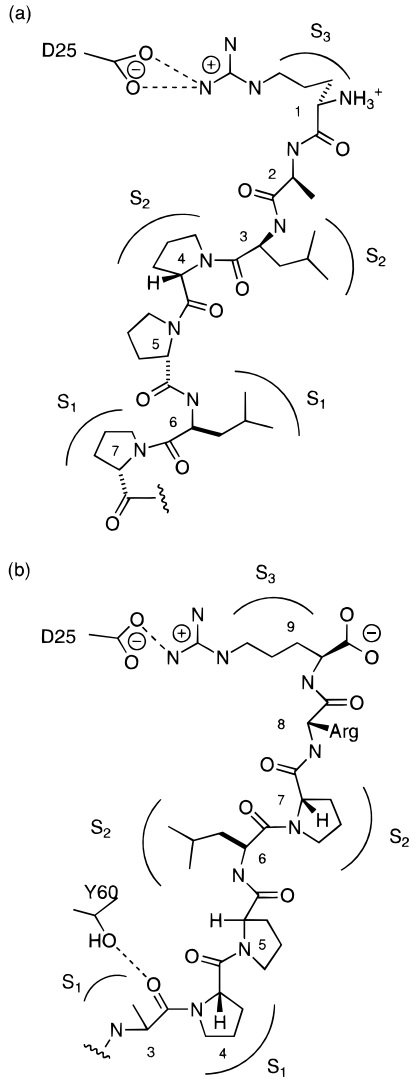


Figure XL*17. Protein–ligand interactions of class I (a) and class II (b) ligands with Src SH3 domain protein.

Two flexible loops connecting the β strands in the SH3 domain, named the n-Src and RT loops, are common to all SH3 domains, and are the primary determinant of ligand specificity for these SH3 domain proteins.[519] This region of the protein is adjacent to the S_3 binding site. Armed with the knowledge of how class I ligands bind to SH3 domains, and the location of the specificity pocket, a "protein structure-based combinatorial library" was designed and prepared.[520] The library was designed to adopt a class I orientation by attaching the biasing sequence PLPPLP ($K_d > 1$ mM) to the solid support. The structure of the class I peptide SH3 complex suggested that the N-terminal proline nitrogen was positioned to introduce groups into the S_3 specificity binding site. Thus, an encoded combinatorial library[521] of the type Cap-M1-M2-M3-PLPPLP-resin, in which Cap and the M*i*'s were composed of a diverse set of organic monomers, was prepared. Screening this library of ~1.1 million members identified 15 beads containing ligands specific for the Src SH3 domain. Decoding revealed two consensus sequences. Ligands **XL-25** ($K_d = 3.4$ μM) and **XL-26** ($K_d = 11$

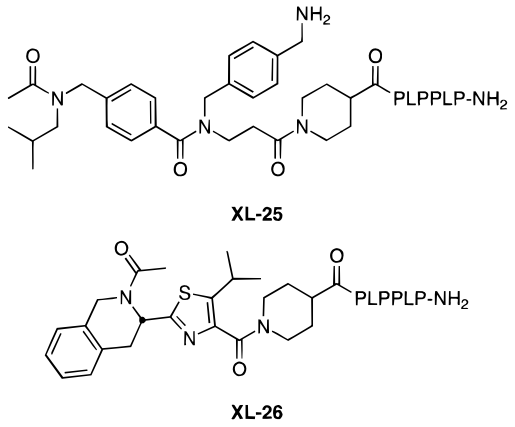


Figure XL*18. Structures of SH3 domain ligands.

μM) are representative members from each class which were subsequently resynthesized on solid phase, cleaved and analyzed for their ability to bind to the Src SH3 domain.

Using NMR spectroscopy the structures of the complexes between both **XL-25** and **XL-26** (Figure XL*18) with the SH3 domain protein were determined (PDB entries: 1NLO and 1NLP).[522] As expected, both ligands bind in the class I orientation, and the nonpeptide elements are located in the S_3 specificity pocket. Thus, the PLPPLP core directs the neighboring nonpeptide groups to the specificity pocket. While peptide ligands make a salt bridge to Asp-99, neither **XL-25** nor **XL-26** make a salt bridge with Asp-99. The nonpeptide portions of both ligands primarily make hydrophobic contacts with the specificity pocket of the protein.

The nonpeptide groups of these ligands thus bind differently to SH3 than do peptide ligands. These results suggest that **XL-25** and **XL-26** would have not been possible to design on the basis of known peptide ligands and provides an illustration of the power of combinatorial chemistry in ligand discovery. In addition to gaining insight into the molecular recognition by SH3 domain proteins, the structures of the complexes between SH3 domains and peptide ligands, **XL-25** and **XL-26** have suggested the design of further biased combinatorial libraries. Peptide ligands form salt bridges with Asp-25 while **XL-25** and **XL-26** do not; therefore, analogs of **XL-25** and **XL-26** which contain basic groups would be expected to be higher affinity ligands and form the basis for further library design.[522]

Thus, this study represents the first example of an iterative cycle of library synthesis and complex structure determination of a few of the highest affinity ligands. These types of studies have great potential for discovering novel ligands, and are a new tool for enhancing our understanding of the molecular recognition of protein−ligand complexes.

IX. Conclusions and Perspectives

A. Biotin–Streptavidin: An Optimized Protein–Ligand Complex

Biotin (**PC-1**) is an enzyme cofactor which functions as a carrier of activated carbon dioxide. Streptavidin is a tetrameric protein that binds biotin with a $K_d = 0.04$ pM. This is one of the tightest protein−

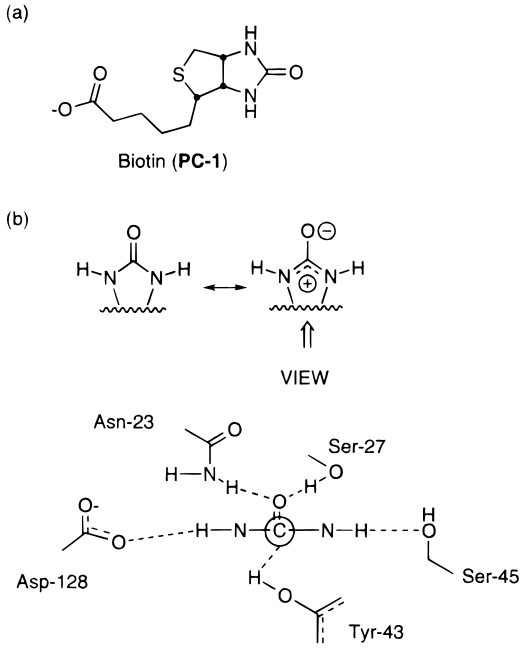


Figure PC*1. (a) Structure of biotin, a $K_d = 4 \times 10^{-14}$ M ligand for streptavidin and (b) summary of electrostatic interactions between biotin and streptavidin. In addition to these interactions there is a very complimentary steric fit between the ligand and the protein.

ligand interactions known. The crystal structure of this complex has been determined by the DuPont Merck group (PDB entry: 1STP).[523]

The structure of this complex reveals a very complimentary steric fit between the ligand and the protein. Most of the solvent accessible surface of biotin is buried within streptavidin. The only solvent exposed group of biotin is the carboxylate. In addition to steric complementarity, there are several complimentary electrostatic interactions. The cyclic urea group is completely buried in the protein and makes extensive hydrogen bonds. (See Figure PC*1.) The carbonyl group makes three hydrogen bonds to protein side chains of Asn-23, Ser-27, and Tyr-43 with an approximately tetrahedral geometry. The side chains of Asn-23 and Ser-27 are properly oriented to bind biotin by additional hydrogen bonds within the protein. The structure of apostreptavidin shows that the groups which bind the urea moiety are preorganized. The two urea NH's make hydrogen bonds to one of the syn lone pairs of Asp-128 and to the alcohol oxygen of Ser-45 respectively. These highly optimized interactions presumably are responsible for the extremely tight binding of biotin to streptavidin.[523]

The side-chain carboxylate of biotin is somewhat solvent exposed. In the crystal structure the syn-lone pairs interact with a water molecule, while the two anti-lone pairs form hydrogen bonds to the backbone NH of Asn-49 and side chain OH of Ser-88. This structure is consistent with the carboxylate group not being crucial for the high affinity of biotin for avidin. The biotin side chain while free from significant torsional strain does have one gauche turn.

Titrating calorimetric measurements have shown that biotin binding is enthalpically favored ($\Delta G° = -18.3$ kcal mol^{-1}, $\Delta H° = -32.0$ kcal mol^{-1}, and $-T\Delta S = 13.7$ kcal mol^{-1}). The authors speculate that the dominant effect contributing to biotin binding is

enhanced hydrogen bonding made between the ligand and protein, due particularly to stabilization of a biotin resonance form whose tetrahedrally coordinated, sp^3 ureido oxygen makes more and stronger hydrogen bonds than an sp^2 oxygen forms with water in solution. That is, streptavidin provides a binding site for the urea in biotin whose interactions stabilize a resonance form of biotin which makes better hydrogen bonds with the protein than are possible with biotin in water.[524] Another interpretation of the calorimetric data is that there is a tightening of much of the protein structure so that the motion of many of the residues becomes more restrictive (giving rise to adverse entropy) and the binding between them becomes stronger (giving rise to the favorable enthalpy). The large changes in exothermicity and large adverse entropies may arise from numerous small changes in the protein. These changes may be too small to be noticed as a distinct structural change.[525] The structure of this complex demonstrates that extremely tight binding is possible with a small ligand which has highly optimized steric and electronic interactions with a protein.

B. Structural Water Molecules

Many proteins have tightly bound water molecules associated with them. For a water to be tightly bound to a protein, the protein−water interaction enthalpy must be greater than the entropic cost of removing a water molecule from bulk solvent and restricting it to a local space on the protein.[40,41] Therefore, inhibitor design strategies which target displacement of these water molecules by incorporating elements of the water molecule within the inhibitor, have great potential.

Protease enzymes can be broadly classified into two generic classes, those which use an enzyme-bound water molecule as the nucleophile (aspartic and zinc metalloproteinases) and those which use a nucleophilic atom of an amino acid side chain (serine, cysteine, and threonine proteases). Most potent inhibitors of aspartic and zinc metallo proteases displace the catalytic water molecule by making favorable interactions with the water binding site. These types of inhibitors should be classified as multisubstrate inhibitors[152] since they bind to both the peptide binding site and the water binding site. These inhibitors take advantage of the ability of the enzyme to bind two substrates, water and polypeptide, simultaneously. Thus, significant additional binding affinity ($\geq$2 kcal mol^{-1}) is available to an inhibitor which binds to the water binding site. If these multisubstrate inhibitors have the appropriate geometry and electrostatic character they may act as mimics of the transition state and achieve greater complementarity to the enzyme. In contrast, serine and cysteine proteases do not tightly bind a substrate water molecule and the multisubstrate inhibition strategy is not an option. Many potent inhibitors of serine and cysteine proteases form covalent adducts with the nucleophilic atom of these proteins to mimic the transition state for peptide hydrolysis and achieve maximum complementarity to the enzyme.

Enzyme-bound water molecules which are not enzyme substrates may also be targeted for displacement by an inhibitor. HIVp protease inhibitors such as cyclic urea **DMP450**[107,109] and the pyrones **AP-32**[119] and **AP-33**[118] incorporate elements of a common structural water ("flap" water) found in most acyclic HIVp inhibitor complexes. As a result, these inhibitors in complex with HIVp do not trap a "flap" water from bulk solvent and gain an entropic advantage over inhibitors which trap this water. The ability of these types of inhibitors to replace the flap water has allowed for a substantial reduction of inhibitor size without sacrificing potent enzyme inhibition. These studies clearly demonstrate that inhibitor design strategies which target displacement of an enzyme-bound water molecule have the potential to yield high affinity ligands. In addition to being a target for displacement, structural water molecules can be considered as a part of the protein. Most HIVp inhibitors make hydrogen bonds to, rather than displace, the "flap" water.

C. Optimization of Electrostatic Interactions

Electrostatic interactions are long range in nature and are also dependent upon the dielectric of medium through which they interact.[196] Many of the examples previously discussed have demonstrated that optimizing electrostatic interactions can be very productive for achieving high affinity ligands.

Structural studies have demonstrated that obtaining optimal electrostatic interactions may be due to rather subtle factors. The study by the DuPont Merck group on thrombin inhibitors related to **DuP714** is very informative.[190] In that study five inhibitors which have positively charged P_1 groups, to interact with the negatively charged S_1 specificity pocket of the enzyme, were studied. The five ligands differ in their ability to bind to thrombin by 4.5 kcal mol^{-1}. X-ray structures for protein−ligand complexes for each ligand were obtained. From this study it is quite clear that simply placing a positive ligand charge in close vicinity to an enzyme negative charge is not enough to achieve optimal electrostatic interactions. This study clearly illustrates that the optimization of ligand−protein electrostatics can be fine tuned by subtle structural modifications.

Hydrogen bonding plays an important role in the stability of protein−ligand complexes. (See Figure PC*2.) A hydrogen bond is formed by the favorable interactions of two dipoles. A dipole−dipole interaction has a $1/r^3$ distance dependence, in addition, there will be an optimal directionality, or angle dependence, for hydrogen bonds.[196] For small molecule crystal structures the mean H···O distance for hydrogen bonds between secondary amides is 1.869 Å, and the mean N···O distance is 2.85 Å. The hydrogen bonds are close to linear with a mean N−H···O angle of 161°, with a standard deviation of 12°. A distinct preference for N−H···O=C hydrogen bonds to be in or near the direction of the carbonyl oxygen atom lone pairs was also observed.[526] If a hydrogen bond is stretched by 0.3 Å or bent by 28°, the energy of the interaction will decrease by approximately 10−15%.[164,527−530] The energetic optimum arrangement occurs when the two dipole moments are colinear.[196] An exhaustive compilation of hydrogen patterns in 15 highly refined protein structures has been published.[531] A carbonyl oxygen, with two lone pairs, is optimally satisfied when it accepts two different

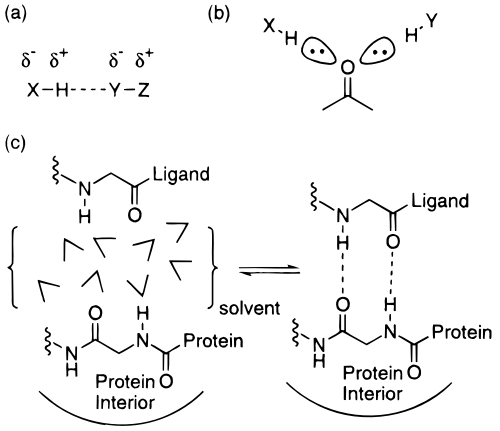
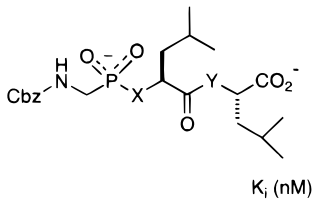


Figure PC*2. (a) Hydrogen bonding is an electrostatic interaction between two dipoles. (b) Hydrogen bonds involving a carbonyl oxygen prefer to be in or near the direction of the carbonyl oxygen atom lone pairs (c) Hydrogen bond formation in aqueous solution is an equilibrium process which is balanced by solvation effects.



		K_i (nM)
PC-2	X = Y = NH	9.1
PC-3	X = O; Y = NH	9000.
PC-4	X = CH_2; Y=NH	10.6
PC-5	X = NH; Y = O	676.

Figure PC*3. Structures and inhibition data of various thermolysin inhibitors.

hydrogen bonds with C═O···H angles close to 120°. However, hydrogen bonds to carbonyl oxygens with a C═O···H angle close to 180° form the basis for β-sheet formation and are quite favorable. The average N—H···O angle is about 155°, with 90% lying between 140° and 180°. In proteins, almost all groups capable of forming hydrogen bonds do so. Where groups are not explicitly hydrogen bonded they are probably solvated. A logical interpretation of this data is that since unsatisfied hydrogen bond donors and acceptors are rarely seen in proteins it is a very unfavorable situation. A corollary for protein—ligand complexes would be burying potential donors and acceptors without forming a complimentary hydrogen bond would also be highly unfavorable. A discussion of hydrogen-bond geometries for 15 different α-lytic protease inhibitor complexes has been published.[164]

A detailed evaluation of the influence of protein—ligand hydrogen bonds to the stability of complexes of thermolysin with phosphorous containing inhibitors has been reported. In this study four series of inhibitors were prepared and evaluated, representative members of each series are **PC-2**, **PC-3**, **PC-4**, and **PC-5**[532] (Figure PC*3). The structure of the most potent inhibitor, **PC-2**, in complex with thermolysin has been reported (PDB entry: 5TMN, see Figure PC*4).[533] This structure shows, among other interactions, a hydrogen bond between the backbone carbonyl of Ala-113 and the phophonamide NH of **PC-2**, and a hydrogen bond between the side chain

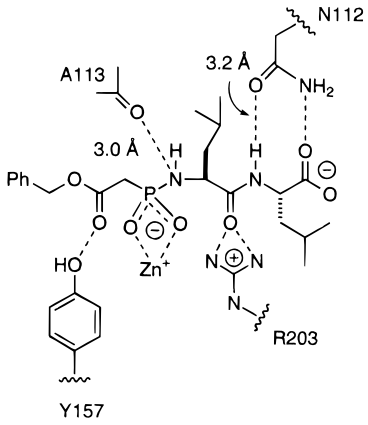


Figure PC*4. Hydrogen bond between thermolysin and **PC-2**.

carbonyl of Asn-112 and the P_2' NH of **PC-2**. The analogous phosphonate **PC-3** was found to be 990-fold less potent (4.1 kcal mol^{-1}) as a thermolysin inhibitor. The structure of the complex between phosphonate **PC-3** and thermolysin has also been reported (PDB entry: 6TMN),[533] and it is nearly identical to the complex between **PC-2** and thermolysin except for the O for NH substitution. The **PC-3** complex has a close repulsive contact (3.1 Å) between the phosphonate ester oxygen and the backbone carbonyl oxygen of Ala-112, no protein conformational change occurs to alleviate this repulsive interaction; this has been referred to as a "forced repulsion".[534] The phosphinate **PC-4** was also prepared and found to be nearly equipotent with **PC-2**, no structure of the **PC-4** complex has been reported. If **PC-4** binds in the same manner as **PC-2** and **PC-3**, then the hydrogen bond between **PC-2** and Ala-112, and the repulsive interaction between **PC-3** and Ala-113, has been replaced with a neutral interaction at the ligand—protein—solvent interface. The observed differences in activity between **PC-2**, **PC-3**, and **PC-4** have been discussed in terms of the balance between active site interactions and solvation. It has been suggested that the phosphonamide is "pulled into the active site by favorable interactions, and the phosphinates are pushed into the active site by unfavorable solvent interactions".[532] An ester (**PC-5**) for amide (**PC-2**) substitution was made and was found to be detrimental for binding. This work, which studies the effect of small changes in inhibitor structure, does not directly determine the energetic contribution of a single hydrogen bond. The net energetic contribution of hydrogen bonds to an equilibrium binding process is balanced by solvation effects. In addition, this work clearly demonstrates the powerful effect of a single repulsive electrostatic interaction on complex stability. An important contribution of hydrogen bonds to complex stability may be the avoidance of the desolvation penalties of polar protein groups upon complex formation. Hydrogen bonds which are formed with optimal geometry will has the most beneficial impact on complex stability.

D. Optimization of VDW Contacts

Attractive VDW interactions occur over a short distance range and attraction decreases as $1/r^6$.[196] As a result, optimization of attractive VDW interactions occurs as the shape of the protein binding site and

the shape of the ligand become more complimentary. Exquisite shape complimentary between a protein and a ligand is rare, even the shape complementarity of the biotin−streptavidin complex is not perfect.[523] However, as long as atoms between the ligand and the protein do not approach too closely, VDW contacts will remain attractive and stabilize the complex. An unfavorable situation will result when the complex leaves a significant gap between the protein and the ligand. Much of the 15-fold difference in affinities (1.6 kcal mol^{-1}) between FKBP12 ligands **AG5120** and **IM-6** is likely due to a gap created in the complex of **AG5120**.[440] In some examples, the protein will alter its shape to maximize protein−ligand VDW contacts. A dramatic illustration of this effect is in the structure of the complex between HIVp and **Ro 31-8959**.[56,70] In this example the protein undergoes a significant conformational change to accommodate the large DIQ group, this results in a high shape complementarity between the DIQ group of the ligand and the protein. Obviously, ligands also alter their conformations to optimize their complementarity to the protein. However, very high energy local conformations, such as eclipsed sp^3-sp^3 bonds, are not common in high affinity ligands.

E. Ligand Conformational Effects

The overall energy of a protein−ligand complex will reflect the net intermolecular energetics of the solvated complex and the intramolecular strain energy of both the protein and the ligand. Several structural studies provide insight into the importance of ligand strain to the overall strength of the complex.

The previously discussed HIVp inhibitor **Ro 31-8959** and its hydroxyl epimer show a 1000-fold difference in affinity toward the enzyme.[56,70] The crystal structures of both complexes have been determined and are nearly identical in their protein−ligand contacts. The major structural difference is that for the weaker inhibitor there is an axial substituent off the DIQ ring system as a result of an inversion of configuration at an sp^3 nitrogen. In this example ligand strain energy is reflected in much poorer enzyme inhibition.

Trifluoromethyl ketones are potent inhibitors of zinc metalloproteases, aspartic proteases and serine proteases[156,535] but are very weak inhibitors of cysteine proteases.[270] Trifluoromethyl ketones in aqueous solution are largely hydrated.[536] For inhibition of zinc metalloproteases and aspartic proteases the hydrated form is the inhibitory species. For cysteine proteases the inhibitory form is presumably the less populated ketone form and the poor enzyme inhibition in this case is likely due to this ligand strain effect. The unhydrated ketone form is also the inhibitory species for serine proteases and trifluoromethyl ketones can be extremely potent inhibitors of this enzyme class. Structural studies have shown that the covalent adducts between TFK inhibitors and serine proteases result in a highly favorable protein−ligand electrostatic interaction.[159−162] Thus, despite a very unfavorable ligand hydration equilibrium, exquisite protein−ligand electrostatic complementarity dominates the overall equilibrium.

Another contributor to favorable ligand binding is ligand conformational preorganization. The cyclic urea inhibitors of HIVp are conformationally constrained ligands.[107] In addition to being constrained inhibitors these compounds also make a unique interaction with the flap of HIVp and prevent the trapping of the "flap" water. Therefore, it is difficult to assess the relative importance of ligand conformational preorganization. Comparison of cyclic urea **AP-28** and the corresponding seco analog **AP-29** indicates a 4.8 kcal mol^{-1} difference in binding energy.[109] However, the bound conformation of **AP-29** has not been determined, and if it bound identically to **AP-28** there would be a severe VDW repulsion between the two carbons which form the ring in the cyclic compound. This example indicates that it is very difficult to obtain acyclic control compounds.

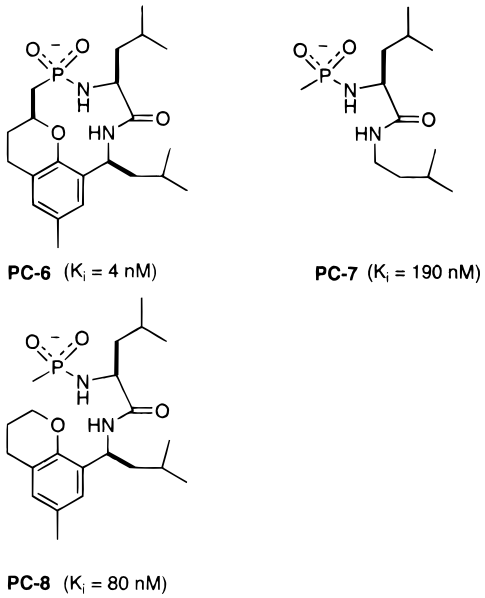


Figure PC*5. Structures of thermolysin inhibitors.

Another example of a conformationally constrained ligand is **PC-6** (Figure PC*5), a thermolysin inhibitor, which is a cyclic analog of **PC-2**.[537] The cyclic ligand **PC-6** sacrifices a hydrogen-bonding interaction between the terminal carboxylate of **PC-2** and the side chain of Asn-113. Despite the loss of that hydrogen bond **PC-6** shows a 2-fold increase in potency relative to **PC-2**. The cyclic inhibitor **PC-6** is also 20-fold more potent than **PC-8** and 50-fold more potent than **PC-7**; both **PC-7** and **PC-8** were prepared as acyclic control compounds. A cocrystal structure of **PC-6** bound to thermolysin revealed that the common features of **PC-6** and **PC-2** bind similarly. Small changes in the enzyme structure were observed to accommodate inhibitor **PC-6** and suggested that the chroman bridging group of **PC-6** was not optimal. A cocrystal structure of the acyclic compound **PC-8** bound to thermolysin revealed an different orientation of the chroman ring than was found in **PC-6**; therefore, it was experimentally verified that **PC-8** was not an appropriate acyclic control compound. The 50-fold increase in affinity of **PC-6**, relative to **PC-7**, is a consequence of both the cyclic conformational constraint and the contacts of the chroman ring of **PC-6** with the enzyme.

A simple, yet elegant, example of biasing a bioactive conformation is the HCA inhibitor **XL-9**.[13] The structure of the HCA complex with inhibitor **XL-6** showed that the ligand bound to the protein in a

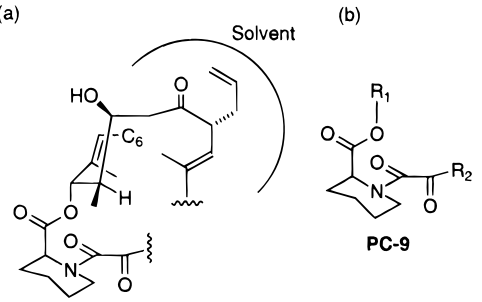


Figure PC*6. (a) **FK506** in its complex with FKBP12 exposes a significant hydrophobic surface into solvent. (b) **PC-9** is an acyclic FKBP12 ligand.

relatively high energy conformation with an axial substituent on a six membered ring. The methyl group of **XL-9** was introduced to reduce the energetic penalty of having an axial substituent on the ring. The disubstituted ring of **XL-9** is forced to have one axial substituent, either the methyl group or the aminoethyl group; as a result the conformational energy of the bound conformer becomes closer to that of the global minimum and reducing the ligand strain penalty upon complex formation. Compound **XL-9** is 2.2 times more potent than the monosubstituted compound **XL-11**.

While the optimal situation for a protein−ligand complex would involve a rigid ligand interacting with the protein in a strain free manner, it appears that this is not a necessary criteria for tight binding. In fact, biotin has a side chain with five rotatable bonds and one them adopts a gauche conformation in its complex with streptavidin. However, factors which reduce overall ligand flexibility and preorganize the ligand in solution for binding to a protein will be beneficial and result in a tighter protein−ligand complex.

While the introduction of appropriate conformation constraints into a ligand can have a beneficial effect of binding, the introduction of constraining elements can also have a deleterious effect on binding. The constrained HIVp inhibitor **AP-38** is 2 orders of magnitude less potent than the unconstrained analog **AP-37**. Structural data indicates that this is due to the constraining elements of **AP-38** making unfavorable steric interactions with the protein.[122] The macrocyclic HIVp inhibitor **AP-39** is 40-fold less potent than the acyclic inhibitor **JG-365**. Structural evidence shows that the common feature of each inhibitor binds identically.[124] It is likely that the reduced affinity for the constrained inhibitor is due to a conformational effect which makes the bound conformer high in energy.

F. Hydrophobic Effects

Literally translated hydrophobic means "fear of water", many of the examples in this review show that this is a very accurate definition. The most dramatic examples of hydrocarbon groups avoiding contact with water comes from studies on acyclic FKBP12 ligands which incorporate FK506-like effector domains (Figure PC*6). FK506, in part because of its macrocyclic nature, exposes a significant amount of hydrophobic surface area into solvent in its complex with FKBP12.[425] Simple acyclic compounds related to **PC-9**, such as **IM-3** (**PC-9**; R_1 = Et, R_2 = CMe_2Et) and **IM-4** (**PC-9**; R_1 = Bn, R_2 = CMe_2Et) are submicromolar ligands for FKBP12. The structure of the **IM-4** complex with FKBP12 shows that it makes protein−ligand interactions similar to FK506 and that these interactions provide for submicromolar affinity.[434] Compounds adorned with appropriate R_1 and R_2 groups, such as **IM-17** and **IM-11** can be low, or sub, nanomolar ligands. However, with certain combinations of R_1 and R_2, such as **AG5163** and **AG5390**, FKBP12 binding affinity is lost.[450]

Analogs **AG5163** and **AG5390** are compounds which have the opportunity to make many of the beneficial interactions made by the submicromolar ligands **IM-3**, **IM-4**, **IM-5**, and **IM-6** and then, like **FK506**, expose a hydrophobic portion of the ligand into solvent. The lack of FKBP12 affinity for these compounds demonstrates that this is not a favorable equilibrium process. A clue as to why these compounds are not FKBP12 ligands comes from the two cocrystal structures of **AG5164**, a close analog of **AG5163**, in complex with FKBP12.[450] The flexible hydrophobic side chains of **AG5163** are not significantly solvent exposed in these complexes, but rather they make nonspecific VDW contacts with the protein surface. That is, the hydrophobic groups of **AG5163** undergo hydrophobic collapse onto the protein surface. The lack of affinity for **AG5163** and **AG5390** may be due to the inability of their hydrophobic groups to find a complimentary patch of protein surface on which to undergo hydrophobic collapse. Another explanation may be that these ligands are flexible, and in solution adopt a number of conformations in which their hydrophobic groups are clustered together to minimize solvent exposure; these hydrophobically collapsed ligand conformations may not be complimentary to FKBP12. Quite possibly, it is a combination of both. While the precise reason for the lack of affinity of **AG5163** and **AG5390** for FKBP12 is not known, it is quite clear that a bound conformation which would expose hydrophobic surface area into solvent is not favorable.

In addition to the **AG5164**−FKBP12 complex, ligand−protein hydrophobic collapse has been observed in the FKBP12 complexes of ligands **IM-9**, **IM-11**, and **IM-31**.[451] Clearly, exposure of hydrophobic surface area in solvent in a protein ligand complex is not an energetically favorable process. The natural products FK506 and rapamycin, and the unnatural ligands **IM-32**[453] and **AG5397_B**,[455] exposure significant ligand hydrophobic surface area into solvent. This occurs only because no other option is available.

It is well known that burying ligand hydrophobic surface area in hydrophobic sites on a protein provides an important driving force in protein−ligand complex formation.[452] In these cases both ligand and protein hydrophobic surfaces are removed from solvent upon complex formation. This is obviously a favorable process. The HIVp active site is quite hydrophobic and as a result most inhibitors have hydrophobic groups and make productive use of ligand−protein hydrophobic contacts. Compound **AG1284**, a potent (K_i = 8 nM) HIVp inhibitor, derives much of its binding affinity from hydrophobic contacts.[64−66] Comparison of the bound conformations of the HIVp inhibitors **L-700,417** and **AP-23** shows

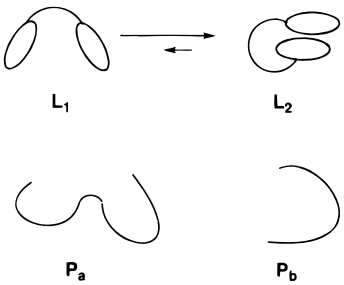


Figure PC*7. Cartoon of productive and hydrophobic collapse. In aqueous solution hydrophobic groups will tend to cluster together to minimize exposed surface area. $\mathbf{L_2}$ represents a favored, hydrophobically collapsed conformation of a flexible ligand. Binding to protein $\mathbf{P_b}$, with a complimentary hydrophobic pocket will be favored by ligand hydrophobic collapse (productive hydrophobic collapse). Binding to protein $\mathbf{P_a}$, with a hydrophobic pocket which is complementary to an extended ligand conformation $\mathbf{L_1}$, will be disfavored by the ligand solution equilibrium (unproductive hydrophobic collapse).

that they differ slightly in the conformation of their common heptanediamide central units. The major consequence of this different conformation is that the cyclohexyl groups in **AP-23** and the phenyl groups of the indanes of **L-700,417** overlap.[96] This suggests that burying complementary hydrophobic surfaces is a major driving force in complex formation.

Ligand hydrophobic collapse is the clustering of hydrophobic groups of a ligand in solution to minimize exposed hydrophobic surface area.[125] A well-described example is the solution conformation of cyclosporin A.[125,538] Another well-described example is the thrombin inhibitor **MD-805**.[125] The bound conformation of **MD-805** with both thrombin and trypsin reveals the stacking of two hydrophobic groups. The high potency of **MD-805** as a thrombin inhibitor suggests that this process, the clustering of two ligand groups, may also occur in solution and induce a bioactive conformation. If this stacking does occur in solution it would be an example of productive hydrophobic collapse. A small molecule crystal structure (crystallized from methanol) of **MD-805** also shows this stacking phenomenon. The structures of the FKBP12 complexes with potent ligands **IM-9** and **IM-11** shows a face-to-face stacking interaction of two ligand aromatic groups.[435,436] It is possible that these ligand stacking interactions may also occur in solution and induce a bioactive conformation. In the complex between HIVp and the potent inhibitor **Ro 31-8959**, the P_1 and P_3 aromatic groups form an edge-to-face interaction in the complex.[70] In the complex between **AG1343** and HIVp the P_1 and P_2 aromatic groups stack upon each other. It is possible that similar hydrophobic interactions also occur in solution and induce local conformations which are productive for tight binding. While in some cases hydrophobic collapse of a ligand in solution can help induce a bioactive conformation in solution, in other cases it may also prevent the formation of solution conformations which are productive for binding. (See Figure PC*7.) The lack of affinity of **AG5163** and **AG5390** for FKBP12 may be due, in part, due to unproductive hydrophobic collapse.

The cyclic nature of the small molecule HIVp inhibitors such as the cyclic urea **DMP323** and the 4-hydroxypyran-2-one **AP-33** likely results in the stability of extended conformations of these compounds in aqueous solution.[118] The relatively high potency of these compounds may be partly due to the prevention of unproductive hydrophobic collapse that can be detrimental to high-affinity binding.

The structures of protein−ligand complexes provide much insight into molecular recognition processes that occur in aqueous solution. Interactions between hydrophobic groups and solvent are highly unfavorable and is some cases can be advantageous and result in high affinity binding, or in other cases can be have extremely deleterious effects of binding phenomenon.

G. Final Comments

The molecular recognition of protein−ligand complexes is highly complex, and our understanding of these equilibrium processes still remains rather elementary. This review has presented a number of examples, some in a detailed fashion, to provide the reader with analogies for future drug-design endeavors. From an experimental point of view, it is essential that some version of an itterative cycle[12] be used in future drug-design efforts.

X. Acknowledgments

We thank Drs. P. Dragovich and B. Burke for critical reading of sections of this manuscript and useful comments, and Drs. K. Appelt and R. Almassy (Agouron), Dr. M. Browner (Roche Bioscience), Dr. S. Fesik (Abbott) and Prof. J. Clardy (Cornell University) for providing crystallographic and NMR coordinates prior to publication.

XI. References

(1) Blundell, T. L. *Nature* **1996**, *384* (SUPP. 7), 23−26.
(2) Petsko, G. A. *Nature* **1996**, *384* (SUPP. 7), 7−9.
(3) Broach, J. R.; Thorner, J. *Nature* **1996**, *384* (SUPP. 7), 14−16.
(4) Bohn, H.-J.; Klebe, G. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2588−2614.
(5) Ringe, D. *Curr. Opin. Struct. Biol.* **1995**, *5*, 825−829.
(6) Karplus, P. A.; Faerman, C. *Curr. Biol.* **1994**, *4*, 770−776.
(7) Searle, M. S.; Williams, D. H. *J. Am. Chem. Soc.* **1992**, *114*, 10690−10697.
(8) Kollman, P. A. *Acc. Chem. Res.* **1996**, *29*, 461−469.
(9) Chambers, J. L.; Stroud, R. M. *Acta Crystallogr. Sect. B.* **1979**, *35*, 1861−1874.
(10) Wlodawer, A.; Nachman, J.; Gilliland, G. L.; Gallagher, W.; Woodward, C. J. *J. Mol. Biol.* **1987**, *198*, 469−480.
(11) Kossiakoff, A. A.; Randal, M.; Guenot, J.; Eigenbrot, C. *Proteins* **1992**, *14*, 65−74.
(12) Appelt, K.; Bacquet, R. J.; Bartlett, C. A.; Booth, C. L. J.; Freer, S. T.; Fuhry, M. A. M.; Gehring, M. R.; Herrmann, S. M.; Howland, E. F.; Janson, C. A.; Jones, T. R.; Kan, C.-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. *J. Med. Chem.* **1991**, *34*, 1925−1934.
(13) Greer, J.; Erickson, J. W.; Baldwin, J. J.; Varney, M. D. *J. Med. Chem.* **1994**, *37*, 1035−1054.
(14) Schreiber, S. L. *Chem. Eng. News.* **1992**, *70* (Oct 26), 22−32.
(15) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasmumi, M. *J. Mol. Biol.* **1977**, *112*, 535−542.
(16) Kahne, D.; Still, W. C. *J. Am. Chem. Soc.* **1988**, *111*, 7529−7534.
(17) Martin, M. T.; Angeles, T. S.; Sugasawara, R.; Aman, N. I.; Napper, A. D.; Darsley, M. J.; Sanchez, R. I.; Booth, P.; Titmas, R. C. *J. Am. Chem. Soc.* **1994**, *116*, 6508−6512.
(18) Radzicka, A.; Wolfenden, R. *J. Am. Chem. Soc.* **1996**, *118*, 6105−6109.
(19) Radzicka, A.; Wolfenden, R. *Science* **1995**, *267*, 90−93.
(20) Pauling, L. *Nature* **1948**, *161*, 707.
(21) Wolfenden, R.; Kati, W. M. *Acc. Chem. Res.* **1991**, *24*, 209−215.

(22) Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157−162.
(23) Lien, E. J.; Gao, H.; Lien, L. L. *Prog. Drug Res.* **1994**, *43*, 43−86.
(24) Vella, S. *AIDS* **1994**, *8* (Suppl. 3), S25−29.
(25) Kitchen, V. S.; Skinner, C.; Ariyoshi, K.; Lane, E. A.; Duncan, I. B.; Burckhardt, J.; Burger, H. U.; Bragman, K.; Pinching, A. J.; Weber, J. N. *Lancet* **1995**, *345*, 952−955.
(26) Baldwin, E. T.; Bhat, T. N.; Gulnik, S.; Hosur, M. V.; Sowder, R. C.; Cachau, R. E.; Collins, J.; Silva, A. M.; Erickson, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6796−6800.
(27) Hyland, L. J.; Tomaszek, T. A., Jr.; Meek, T. D. *Biochemistry* **1991**, *30*, 8454−63.
(28) Fruton, J. S. *Adv. Enzymol. Relat. Areas. Mol. Biol.* **1976**, *44*, 1−36.
(29) Dunn, B. M. *Adv. Detailed React. Mech.* **1992**, *2*, 213−241.
(30) Suguna, K.; Padlan, E. A.; Smith, C. W.; Carlson, W. D.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7009−7013.
(31) James, M. N. G.; Sielecki, A. R.; Hayakawa, K.; Gelb, M. H. *Biochemistry* **1992**, *31*, 3872−3886.
(32) Umezawa, H.; Agoyagi, T.; Morishima, H.; Matzusaki, M.; Hamada, H.; Takeuchi, T. *J. Antibiot.* **1970**, *23*, 259−262.
(33) Rich, D. H. *J. Med. Chem.* **1985**, *28*, 263−273.
(34) Bott, R.; Subramanian, E.; Davies, D. R. *Biochemistry* **1982**, *21*, 6956−6962.
(35) Holladay, M. W.; Salituro, F. G.; Rich, D. H. *J. Med. Chem.* **1987**, *30*, 374−383.
(36) Salituro, F. G.; Agarwal, N.; Hofman, T.; Rich, D. H. *J. Med. Chem.* **1987**, *30*, 286−295.
(37) Rich, D. H.; Sun, E. T. O.; Ulm, E. *J. Med. Chem.* **1980**, *23*, 27−33.
(38) Rich, D. H.; Sun, E.; Singh, J. *Biochem. Biophys. Res. Commun.* **1977**, *74*, 762−764.
(39) Rich, D. H.; Bernatowicz, M. S.; Schmidt, P. G. *J. Am. Chem. Soc.* **1982**, *104*, 3535−3536.
(40) Dunitz, J. D. *Science* **1994**, *264*, 670.
(41) Bryan, W. P. *Science* **1994**, *266*, 1726.
(42) Dunitz, J. D. *Chem. Biol.* **1995**, *2*, 709−712.
(43) Bartlett, P. A.; Kezer, W. B. *J. Am. Chem. Soc.* **1984**, *106*, 4282−4283.
(44) Bartlett, P. A.; Hanson, J. E.; Giannousis, P. P. *J. Org. Chem.* **1990**, *55*, 8268−6274.
(45) Fraser, M. E.; Stryndka, N. C. J.; Bartlett, P. A.; Hanson, J. E.; James, M. N. G. *Biochemistry* **1992**, *31*, 5201−5214.
(46) Bartlett, P. A.; Giangiordano, M. A. *J. Org. Chem.* **1996**, *61*, 3433−3438.
(47) Boger, J.; Lohr, N. S.; Ulm, E. H.; Poe, M.; Blaine, E. H.; Fanelli, G. M.; Lin, T.-Y.; Payne, L. S.; Schorn, T. W.; LaMont, B. I.; Vassil, T. C.; Stabilito, I. I.; Veber, D. F.; Rich, D. H.; Bopari, A. S. *Nature* **1983**, *303*, 81−84.
(48) Katoh, I.; Yasunga, T.; Ikawa, Y.; Yoshinaka, Y. *Nature* **1987**, *329*, 654−656.
(49) Fitzgerald, P. M. D.; McKeever, B. M.; VanMiddlesworth, J. F.; Springer, J. P.; Heimbach, J. C.; Leu, C.-T.; Herber, W. K.; Dixon, R. A. F.; Darke, P. L. *J. Biol. Chem.* **1990**, *265*, 14209−14219.
(50) (a) Greenlee, W. J.; Siegl, P. K. S. *Annu. Rep. Med. Chem.* **1992**, *27*, 59−68. (b) Abdel-Meguid, S. S. *Med. Chem. Rev.* **1993**, *13*, 731−778.
(51) Dhanaraj, V.; Dealwis, C. G.; Frazao, C.; Badasso, M.; Sibanda, B. L.; Tickle, I. J.; Cooper, J. B.; Driessen, H. P.; Newman, M.; Aguilar, C.; Wood, S. P.; Blundell, T. L.; Hobart, P. M.; Geoghegan, K. F.; Ammirati, M. J.; Danley, D. E.; O'Connor, B. A.; Hoover, D. J. *Nature* **1992**, *357*.
(52) Navia, M. A.; Fitzgerald, P. M.; McKeever, B. M.; Leu, C. T.; Heimbach, J. C.; Herber, W. K.; Sigal, I. S.; Darke, P. L.; Springer, J. P. *Nature* **1989**, *337*, 615−620.
(53) Wlodawer, A.; Miller, M.; Jaskolski, M.; Sathyanarayana, B. K.; Baldwin, E.; Weber, I. T.; Selk, L. M.; Clawson, L.; Schneider, J.; S. B. Kent, S. B. *Science* **1989**, *245*, 616−621.
(54) Huff, J. R. *J. Med. Chem.* **1991**, *34*, 2305−2314.
(55) Tomasselli, A. G.; Howe, W. J.; Sawyer, T. K.; Wlodawer, A.; Heinrikson, R. L. *Chim. Oggi* **1991**, *9*, 6−27.
(56) Appelt, K. *Perspect. Drug Discovery Des.* **1993**, *1*, 23−48.
(57) Wlodawer, A.; Erickson, J. W. *Annu. Rev. Biochem.* **1993**, *62*, 543−585.
(58) Jaskolski, M.; Tomasselli, A. G.; Sawyer, T. K.; Staples, D. G.; Heinrikson, R. L.; Schneider, J.; Kent, S. B. H.; Wlodawer, A. *Biochemistry* **1991**, *30*, 1600−1609.
(59) Dreyer, G. B.; Lambert, D. M.; Meek, T. D.; Carr, T. J.; Tomaszek, T. A., Jr.; Fernandez, A. V.; Bartus, H.; Cacciavillani, E.; Hassell, A. M.; Minnich, M.; Petteway, S. R., Jr.; Metcalf, B. W. *Biochemistry* **1992**, *31*, 6646−6659.
(60) Abdel-Meguid, S. S.; Metcalf, B. W.; Carr, T. J.; Demarsh, P.; DesJarlais, R. L.; Fisher, S.; Green, D. W.; Ivanoff, L.; Lambert, D. M.; Murthy, K. H. M.; Petteway, S. R. J.; Pitts, W. J.; Tomaszek, T. A. J.; Winborne, E.; Zhao, B.; Dreyer, G. B.; Meek, T. D. *Biochemistry* **1994**, *33*, 11671−11677.
(61) Thompson, S. K.; Murthy, K. H. M.; Zhao, B.; Winborne, E.; Green, D. W.; Fisher, S. M.; DesJarlais, R. L.; Tomaszek, T. A., Jr.; Meek, T. D.; Gleason, J. G.; Abdel-Meguid, S. S. *J. Med. Chem.* **1994**, *37*, 3100−3107.
(62) Newlander, K. A.; Callahan, J. F.; Moore, M. L.; Tomaszek, T. A., Jr.; Huffman, W. F. *J. Med. Chem.* **1993**, *36*, 2321−31.
(63) Hoog, S. S.; Zhao, B.; Winborne, E.; Fisher, S.; Green, D. W.; DesJarlais, R. L.; Newlander, K. A.; Callahan, J. F.; Moore, M. L.; Huffman, W. F.; Abdel-Meguid, S. S. *J. Med. Chem.* **1995**, *38*, 3246−3252.
(64) Reich, S. H.; Melnick, M.; Davies, J. F., II; Appelt, K.; Lewis, K. K.; Fuhry, M. A.; Pino, M.; Trippe, A. J.; Nguyen, D.; Dawson, H.; Wu, B. W.; Musick, L.; Kosa, M.; Kahil, D.; Webber, S.; Gehlhaar, D. K.; Andrada, D.; Shetty, B. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3298−3302.
(65) Reich, S. H.; Melnick, M.; Pino, M.; Fuhry, M. A.; Trippe, A. J.; Appelt, K.; Davies, J. F., II; Wu, B. W.; Musick, L. *J. Med. Chem.* **1996**, *39*, 2781−2794.
(66) Melnick, M.; Reich, S. H.; Lewis, K. K.; Mitchell, L. J., Jr.; Nguyen, D.; Trippe, A. J.; Dawson, H.; Davies, J. F., II; Appelt, K.; Wu, B. W.; Musick, L.; Gehlhaar, D. K.; Webber, S.; Shetty, B.; Kosa, M.; Kahill, D.; Andrada, D. J. *Med. Chem.* **1996**, *39*, 2795−2811.
(67) Getman, D. P.; DeCrescenzo, G. A.; Heintz, R. M.; Reed, K. L.; Talley, J. J.; Bryant, M. L.; Clare, M.; Houseman, K. A.; Marr, J. J.; Mueller, R. A.; Vazquez, M. L.; Shieh, H. S.; Stallings, W. C.; Stegeman, R. A. *J. Med. Chem.* **1993**, *36*, 288−291.
(68) Rich, D. H.; Green, J.; Toth, M. V.; Marshall, G. R.; Kent, S. B. H. *J. Med. Chem.* **1990**, *33*, 1285−1288.
(69) Roberts, N. A.; Martin, J. A.; Kinchington, D.; Broadhurst, A. V.; Craig, J. C.; Duncan, I. B.; Galpin, S. A.; Handa, B. K.; Kay, J.; Krohn, A.; Lambert, R. W.; Merrett, J. H.; Mills, J. S.; Parkes, K. E. B.; Redshaw, S.; Ritchie, A. J.; Taylor, D. L.; Thomas, G. J.; Machin, P. J. *Science* **1990**, *248*, 358−361.
(70) Krohn, A.; Redshaw, S.; Ritchie, J. C.; Graves, B. J.; Hatada, M. H. *J. Med. Chem.* **1991**, *34*, 3340−2.
(71) Rich, D. H.; Sun, C. Q.; Prasad, J. V. N. V.; Pathiasseril, A.; Toth, M. V.; Marshall, G. R.; Clare, M.; Mueller, R. A.; Houseman, K. *J. Med. Chem.* **1991**, *34*, 1222−5.
(72) Tam, T. F.; Carriere, J.; MacDonald, I. D.; Castelhano, A. L.; Pliura, D. H.; Dewdney, N. J.; Thomas, E. M.; Bach, C.; Barnett, J.; Chan, H.; Krantz, A. *J. Med. Chem.* **1992**, *35*, 1318−1320.
(73) Baldwin, E. T.; Bhat, E. T.; Gulnik, S.; Liu, B.; Topol, I. A.; Kiso, Y.; Mimoto, T.; Mitsuya, H.; Erickson, J. W. *Structure* **1995**, *3*, 581−590.
(74) Wang, Y.-X.; Freedberg, D. I.; Yamazaki, T.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Kiso, Y.; Torchia, D. A. *Biochemistry* **1996**, *35*, 9945−9950.
(75) Slee, D. H.; Laslo, K. L.; Elder, J. H.; Ollman, I. R.; Gustchina, A.; Kervinen, J.; Zdanov, A.; Wlodawer, A.; Wong, C.-H. *J. Am. Chem. Soc.* **1995**, *117*, 11867−11878.
(76) Kaldor, S. W.; Hammond, M.; Dressman, B. A.; Fritz, J. E.; Crowell, T. A.; Hermann, R. A. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1385−90.
(77) Kaldor, S. W.; Dressman, B. A.; Hammond, M.; Appelt, K.; Burgess, J.; Lubbehusen, P. P.; Muesing, M. A.; Hatch, S. D.; Wiskerchen, M. A.; Baxter, A. J. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 721−726.
(78) Kaldor, S. W.; Appelt, K.; Fritz, J.; Hammond, M.; Crowell, T. A.; Baxter, A. J.; Hatch, S. D.; Wiskerchen, M. A.; Muesing, M. A. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 715−720.
(79) Kalish, V. J.; Tatlock, J. H.; Davies, J. F.; Kaldor, S. W.; Dressman, B. A.; Reich, S.; Pino, M.; Nyugen, D.; Appelt, K.; Musick, L.; Wu, B.-w. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 727−732.
(80) Kalish, V.; Kaldor, S.; Shetty, B.; Tatlock, J.; Davies, J.; Hammond, M.; Dressman, B.; Fritz, J.; Appelt, K.; Reich, S., R.; Musick, L.; Wu, B.-w.; Su, K. *Eur. J. Med. Chem.* **1995**, *30*, 201s−214s.
(81) Kim, E. E.; Baker, C. T.; Dwyer, M. D.; Murcko, M. A.; Rao, B. G.; Tung, R. D.; Navia, M. A. *J. Am. Chem. Soc.* **1995**, *117*, 1181−1182.
(82) Ghosh, A. K.; Thompson, W. J.; Fitzgerald, P. M. D.; Culberson, J. C.; Axel, M. G.; McKee, S. P.; Huff, J. R.; Anderson, P. S. *J. Med. Chem.* **1994**, *37*, 2506−2508.
(83) Ghosh, A. K.; Thompson, W. J.; McKee, S. P.; Duong, T. T.; Lyle, T. A.; Chen, J. C.; Darke, P. L.; Zugay, J. A.; Emini, E. A.; Schleif, W. A.; Huff, J. R.; Anderson, P. S. *J. Med. Chem.* **1993**, *36*, 292−294.
(84) Ghosh, A. K.; Lee, H. Y.; Thompson, W. J.; Culberson, C.; Holloway, M. K.; McKee, S. P.; Munson, P. M.; Duong, T. T.; Smith, A. M.; Darke, P. L.; Zugay, J. A.; Emini, E. A.; Schleif, W. A.; Huff, J. R.; Anderson, P. S. *J. Med. Chem.* **1994**, *37*, 1177−1188.
(85) Thompson, W. J.; Ghosh, A. K.; Holloway, M. K.; Lee, H. Y.; Munson, P. M.; Schwering, J. E.; Wai, J.; Darke, P. L.; Zugay, J.; Emini, E. A.; Schleif, W. A.; Huff, J. R.; Anderson, P. S. *J. Am. Chem. Soc.* **1993**, *115*, 801−803.
(86) Blundell, T.; Pearl, L. *Nature* **1989**, *337*, 596−7.
(87) 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. *Science* **1990**, *249*, 527−533.
(88) 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. C.; Saldivar, A.; Rosenbrook, W., Jr.; Plattner, J. J.; Erickson, J. *J. Med. Chem.* **1990**, *33*, 2687−2689.
(89) 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. *J. Am. Chem. Soc.* **1994**, *116*, 847−855.
(90) Kempf, D. J.; Marsh, K. C.; Denissen, J. F.; McDonald, E.; Vasavanonda, S.; Flentge, C. A.; Green, B. E.; Fino, L.; Park, C. H.; Kong, X.-P.; Wideburg, N. E.; Salvidar, A.; Ruiz, L.; Kati, W. M.; Sham, H. L.; Robins, T.; Stewart, K. D.; Hsu, A.; Plattner, J. J.; Leonard, J. M.; Norbeck, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2484−2488.
(91) 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., Jr.; Meek, T. D.; Dreyer, G. B. *Biochemistry* **1993**, *32*, 7972−7980.
(92) Harte, W. E., Jr.; Beveridge, D. L. *J. Am. Chem. Soc.* **1993**, *115*, 3883−3886.
(93) Bone, R.; Vacca, J. P.; Anderson, P. S.; Holloway, M. K. *J. Am. Chem. Soc* **1991**, *113*, 9382−9384.
(94) Babine, R. E.; Zhang, N.; Jurgens, A. R.; Schow, S. R.; Desai, P. R.; James, J. C.; Semmelhack, M. F. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 541−546.
(95) Babine, R. E.; Zhang, N.; Schow, S. R.; Jirousek, M. R.; Johnson, B. D.; Kerwar, S. S.; Desai, P. R.; Byrn, R. A.; Hastings, R. C.; Wick, M. M. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1589−1594.
(96) Babine, R. E.; Zhang, N.; Schow, S. R.; Xu, Z.; Byrn, R. A.; Hastings, R. C.; Semmelhack, M. F.; Wick, M. M.; Kerwar, S. S. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 583−588.
(97) Trova, M. P.; Babine, R. E.; Byrn, R. A.; Casscles, W. T., Jr.; Hastings, R. C.; Hsu, G. C.; Jirousek, M. R.; Johnson, B. D.; Kerwar, S. S.; Schow, S. R.; Wissner, A.; Zhang, N.; Wick, M. M. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1595−1600.
(98) Trova, M. P.; Wissner, A.; Casscles, W. T., Jr.; Hsu, G. C. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 903−906.
(99) Vacca, J. P.; Dorsey, B. D.; Schleif, W. A.; Levin, R. B.; McDaniel, S. L.; Darke, P. L.; Zugay, J.; Quintero, J. C.; Blahy, O. M.; Roth, E.; Sardana, V. V.; Schlabach, A. J.; Graham, P. I.; Condra, J. H.; Gotlib, L.; Holloway, M. K.; Lin, J.; Chen, I.-W.; Vastag, K.; Ostovic, D.; Anderson, P. S.; Emini, E. A.; Huff, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4096−4100.
(100) Dorsey, B. D.; Levin, R. B.; McDaniel, S. L.; Vacca, J. P.; Guare, J. P.; Darke, P. L.; Zugay, J. A.; Emini, E. A.; Schleif, W. A.; Quintero, J. C.; Lin, J. H.; Chen, I.-W.; Holloway, M. K.; Fitzgerald, P. M. D.; Axel, M. G.; Ostovic, D.; Anderson, P. S.; Huff, J. R. *J. Med. Chem.* **1994**, *37*, 3443−3451.
(101) Appelt, K.; Babine, R. E.; Schow, S. R. Unpublished results.
(102) Chen, Z.; Li, Y.; Chen, E.; Hall, D.; Darke, P.; Culberson, C.; Shafer, J. A.; Kuo, L. A. *J. Biol. Chem.* **1994**, *269*, 26344−26348.
(103) Wonacott, A.; Cooke, R.; Hayes, F. R.; Hann, M. M.; Jhoti, H.; McMeekin, P.; Mistry, A.; Murray-Rust, P.; Singh, O. M. P.; Weir, M. P. *J. Med. Chem.* **1993**, *36*, 3113−3119.
(104) Jhoti, H.; Singh, O. M. P.; Weir, M. P.; Cooke, R.; Murray-Rust, P.; Wonacott, A. *Biochemistry* **1994**, *33*, 8417−8427.
(105) Barrish, J. C.; Gordon, E.; Alam, M.; Lin, P.-F.; Bisacchi, G. S.; Chen, P.; Cheng, P. T. W.; Fritz, A. W.; Greytok, J. A.; Hermeiser, M. A.; Humphreys, W. G.; Lis, K. A.; Marella, M. A.; Merchant, Z.; Mitt, T.; Morrison, R. A.; Obermeier, M. T.; Pluscec, J.; Skoog, M.; Slusarchyk, W. A.; Spergel, S. H.; Stevenson, J. M.; Sun, Q.-q.; Sundeen, J. E.; Taunk, P.; Tino, J. A.; Warrack, B. M.; Colonno, R. J.; Zahler, R. *J. Med. Chem.* **1994**, *37*, 1758−1768.
(106) Bisacchi, G. S.; Ahmad, S.; Alam, M.; Ashfaq, A.; Barrish, J.; Cheng, P. T. W.; Greytok, J.; Hermsmier, M.; Lin, P. F.; Merchant, Z.; Skoog, M.; Spergal, S.; Zahler, R. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 459−464.
(107) Lam, P. Y. S.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, O. M. J.; Rayner, M. M. *Science* **1994**, *263*, 380−384.
(108) Hodge, C. N.; Aldrich, P. E.; Bacheler, L. T.; Chang, C.-H.; Eyerman, C. J.; Gribb, M. F.; Jackson, D. A.; Jadhav, P. K.; Korant, B.; Lam, P. Y. S.; Maurin, M. B.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Sharpe, T. R.; Shum, L.; Winslow, D. L.; Erickson-Viitanen, S. *Chem. Biol.* **1996**, *3*, 301−314.
(109) Lam, P. Y. S.; Ru, Y.; Jadhav, P. K.; Aldrich, P. E.; DeLucca, G. V.; Eyermann, C. J.; Chang, C.-H.; Emmett, G.; Holler, E. R.; Daneker, W. F.; Li, L.; Confalone, P. N.; McHugh, R. J.; Han, Q.; Li, R.; Markwalder, J. A.; Seitz, S. P.; Sharpe, T. R.; Bacheler, L. T.; Rayner, M. M.; Klabe, R. M.; Shum, L.; Winslow, D. L.; Kornhauser, D. M.; Kackson, D. A.; Erickson-Viitanen, S.; Hodge, C. N. *J. Med. Chem.* **1996**, *39*, 3514−3525.
(110) Yamazaki, T.; Nicholson, L. K.; Torchia, D. A.; Wingfield, P.; Stahl, S. J.; Kaufman, J. D.; Eyermann, C. J.; Hodge, C. N.; Lam, P. Y. S.; Ru, Y.; Jadhav, P. K.; Chang, C.-H.; Weber, P. C. *J. Am. Chem. Soc.* **1994**, *116*, 10791−10792.
(111) Bures, M. G.; Hutchins, C. W.; Maus, M.; Kohlbrenner, W.; Kadam, S.; Erickson, J. W. *Tetrahedron Comput. Methodol.* **1990**, *3*, 673−680.
(112) Chenera, B.; DesJarlais, R. L.; Finkelstein, J. A.; Eggleston, D. S.; Meek, T. D.; Tomaszek, T. A., Jr.; Dreyer, G. B. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 2717−2722.
(113) Randad, R. S.; Pan, W.; Gulnik, S. V.; Burt, S.; Erickson, J. W. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1247−1252.
(114) Peyman, A.; Stahl, W.; Wagner, K.; Ruppert, D.; Budt, K.-H. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2601−2604.
(115) Kim, C. U.; McGee, L. R.; Krawczyk, S. H.; Harwood, E.; Harada, Y.; Swaminathan, S.; Bischofberger, N.; Chen, M. S.; Cherrington, J. M.; Xiong, S. F.; Griffin, L.; Cundy, K. C.; Lee, A.; Yu, B.; Gulnik, S.; Erickson, J. W. *J. Med. Chem.* **1996**, *39*, 3431−3434.
(116) Thaisrivongs, S.; Tomich, P. K.; Watenpaugh, K. D.; Chong, K.-T.; Howe, W. J.; Yang, C.-P.; Strohbach, J. W.; Turner, S. R.; McGrath, J. P.; Bohanon, M. J.; Lynn, J. C.; Mulichak, A. M.; Spinelli, P. A.; Hinshaw, R. R.; Pagano, P. J.; Moon, J. B.; Ruwart, M. J.; Wilkenson, K. F.; Rush, B. D.; Zipp, G. L.; Dalga, R. J.; Schwende, F. J.; Howard, G. M.; Padbury, G. E.; Toth, L. N.; Zhao, Z.; Koeplinger, K. A.; Kakuk, T. J.; Cole, S. L.; Zaya, R. M.; Piper, R. C.; Jeffrey, P. *J. Med. Chem.* **1994**, *37*, 3200−3204.
(117) (a) Vara Prasad, J. V. N.; Para, K. S.; Lunney, E. A.; Ortwine, D. F.; Dunbar, J. B., Jr.; Ferguson, D.; Tummino, P. J.; Hupe, D.; Tait, B. D.; Domagla, J. M.; Humblet, C.; Bhat, T. N.; Liu, B.; Guerin, D. M. A.; Baldwin, E. T.; Erickson, J. W.; Sawyer, T. K. *J. Am. Chem. Soc.* **1994**, *116*, 6989−6990. (b) Lunney, E. A.; Hagen, S. E.; Domagala, J.; Humblet, C.; Tait, B.; Warmus, J.; Wilson, M.; Ferguson, D.; Hupe, D.; Tummino, P. J.; Baldwin, E.; Bhat, T. N.; Liu, B.; Erickson, J. W. *J. Med. Chem.* **1994**, *37*, 2664−2677.
(118) Vara Prasad, J. V. N.; Lunney, E. A.; Ferguson, D.; Tummino, P. J.; Rubin, J. R.; Reyner, E. L.; Stewart, B. H.; Guttendorf, R. J.; Domagala, J. M.; Suvorov, L. I.; Gulnik, S. V.; Topol, I. A.; Bhat, T. N.; Erickson, J. W. *J. Am. Chem. Soc* **1995**, *117*, 11070−11074.
(119) Romines, K. R.; Watenpaugh, K. D.; Howe, W. J.; Tomich, P. K.; Lovasz, K. D.; Morris, J. K.; Janakiraman, M. N.; Lynn, J. C.; Horng, M.-M.; Chong, K.-T.; Hinshaw, R. R.; Dolak, L. A. *J. Med. Chem.* **1995**, *38*, 4463−4473.
(120) Vara Prasad, J. V. N.; Pavlovshy, A.; Para, K. S.; Ellsworth, E. L.; Tummino, P. J.; Nouhan, C.; Ferguson, D. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1133−1138.
(121) Kim, B. M.; Vacca, J. P.; Fitzgerald, P. M. D.; Darke, P. L.; Holloway, M. K.; Guare, J. P.; Hanifin, C. M.; Arford-Bickerstaff, D. J.; Zugay, J. A.; Wai, J. M.; Anderson, P. S.; Huff, J. R. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2199−2204.
(122) Vacca, J. P.; Fitzgerald, P. M. D.; Holloway, M. K.; Hungate, R. W.; Starbuck, K. E.; Chen, L. J.; Darke, P. L.; Anderson, P. S.; Huff, J. R. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 499−504.
(123) Smith, R. A.; Coles, P. J.; Chen, J. J.; Robinson, V. J.; MacDonald, I. D.; Carriere, J.; Krantz, A. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2217−2222.
(124) Abbenante, G.; March, D. R.; Bergman, D. A.; Hunt, P. A.; Garnham, B.; Dancer, R. J.; Martin, J. L.; Fairlie, D. P. *J. Am. Chem. Soc.* **1995**, *117*, 10220−10226.
(125) Wiley, R. A.; Rich, D. H. *Med. Res. Rev.* **1993**, *13*, 327−384.
(126) Fersht, A. *Enzyme Structure and Mechanism*, 2nd ed.; W. H. Freeman and Company: New York, 1985; pp 405−413.
(127) Gibson, W.; Welch, A. R.; Hall, M. R. T. *Perspect. Drug Discovery Des.* **1994**, *2*, 413−426.
(128) Chen, P.; Tsuge, H.; Almassy, R. J.; Gribskov, C. L.; Katoh, S.; Vanderpool, D. L.; Margosiak, S. A.; Pinko, C.; Matthews, D. A.; Kan, C.-C. *Cell* **1996**, *86*, 835−843.
(129) Tong, L.; Qian, C.; Massariol, M.-J.; Bonneau, P. R.; Cordingley, M. G.; Lagace, M. G. *Nature* **1996**, *383*, 272−275.
(130) Qui, X.; Culp, J. S.; DiLella, A. G.; Hellmig, B.; Hoog, S. S.; Janson, C. A.; Smith, W. W.; Abdel-Meguid, S. S. *Nature* **1996**, *383*, 275−279.
(131) Shieh, H.-S.; Kurumbail, R. G.; Stevens, A. M.; Stegeman, R. A.; Sturman, E. J.; Pak, J. Y.; Wittwer, A. J.; Palmier, M. O.; Wiegand, R. C.; Holwerda, B. C.; Stallings, W. C. *Nature* **1996**, *383*, 279−282.
(132) Love, R. A.; Parge, H. E.; Wickersham, J. A.; Hostomsky, Z.; Habuka, N.; Moomaw, E. W.; Adachi, T.; Hostomska, Z. *Cell* **1996**, *87*, 331−342.
(133) Kim, J. L.; Morgenstern, K. A.; Lin, C.; Fox, T.; Dwyer, M. D.; Landro, J. A.; Chambers, S. P.; Markland, W.; Lepre, C. A.; O'Malley, E. T.; Harbeson, S. L.; Rice, C. M.; Murcko, M. A.; Caron, P. R.; Thomson, J. A. *Cell* **1996**, *87*, 343−355.
(134) O'Leary, M. H.; Kluetz, M. D. *J. Am. Chem. Soc.* **1972**, *94*, 3585−3589.
(135) Wright, C. S.; Alden, R. A.; Kraut, J. *Nature* **1969**, *221*, 235.
(136) Bode, W.; Huber, R. *Eur. J. Biochem.* **1992**, *204*, 433−451.
(137) Frigerio, F.; Coda, A.; Pugliese, L., C.; Mengatti, E.; Amiconi, G.; Schnebli, H. P.; Ascenzi, P.; Bolognesi, M. *J. Mol. Biol.* **1992**, *225*, 107−123.
(138) Bode, W.; Papamokos, E.; Musil, D. *Eur. J. Biochem.* **1987**, *166*, 673−692.
(139) Read, R. J.; James, M. N. G. In *Protease Inhibitors*; Barrett, A. J., Salveseb, G., Eds.; Elsevier: Amsterdam, 1986; pp 301−336.
(140) Umezawa, H.; Aoyagi, T.; Morishima, H.; Kunimoto, S.; Matsuzaki, M.; Hamada, M.; Takeuchi, T. *J. Antibiot.* **1970**, *23*, 425−427.
(141) Umezawa, H. *Acta Biol. Med. Ger.* **1977**, *36*, 1899−1915.

(142) Umezawa, H. *Annu. Rev. Microbiol.* **1982**, *36*, 75−99.
(143) Delbaere, L. T. J.; Brayer, G. D. *J. Mol. Biol.* **1985**, *183*, 89.
(144) Kurinov, I. V.; Harrison, R. W. *Protein Sci.* **1996**, *5*, 752−758.
(145) Berglund, G. I.; Smalas, A. O.; Hordvik, A.; Williassen, N. P. *Acta Crystallogr., Sect. D* **1995**, *D51*, 725−730.
(146) Lee, A. Y.; Smitka, T. A.; Bonjouklian, R.; Clardy, J. *Chem. Biol.* **1994**, *1*, 113−117.
(147) Lee, A. Y.; Hagihara, M.; Karmacharya, R.; Albers, M. W.; Schreiber, S. L.; Clardy, J. *J. Am. Chem. Soc.* **1993**, *115*, 12619−12620.
(148) Maryanoff, B. E.; Qui, X.; Padmanabhan, K. P.; Tulinsky, A.; Almond, H. R.; Andrade-Gorgon, P.; Greco, M. N.; Kauffman, J. A.; Nicolaou, K. C.; Liu, A.; Btungs, P. H.; Fusetani, N. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8048−8052.
(149) Ganesh, V.; Lee, A. Y.; Clardy, J.; Tulinsky, A. *Protein Sci.* **1996**, *5*, 825−835.
(150) Banner, D.; Ackermann, J.; Gast, A.; Gubernator, K.; Hadvary, P.; Hilpert, K.; Labler, L.; Mueller, K.; Schmid, G.; Tschopp, T.; Waterbeemd, H. v. d.; Wirz, B. In *Perspect. Med. Chem.*; Testa, B., Ed.; Verlag Helvetica Chim. Acta: Basel, Switzerland, 1993.
(151) Nienaber, V. L.; Amparo, E. C. *J. Am. Chem. Soc.* **1996**, *118*, 6807−6810.
(152) Krantz, A. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 1327−1334.
(153) Thompson, R. C. *Biochemistry* **1973**, *12*, 47.
(154) Shah, D. O.; Lai, K.; Gornstein, D. G. *J. Am. Chem. Soc.* **1984**, *106*, 4272−4273.
(155) James, M. N. G.; Sielecki, A. R.; Brayer, G. D.; Delbaere, L. T. J.; Bauer, C. A. *J. Mol. Biol.* **1980**, *144*, 43−88.
(156) Imperiali, B.; Abeles, R. H. *Biochemistry* **1986**, *25*, 3760−3767.
(157) Brady, K.; Wei, A.; Ringe, D.; Abeles, R. H. *Biochemistry* **1990**, *29*, 7600−7607.
(158) Liang, T.-C.; Abeles, R. H. *Arch. Biochem. Biophys.* **1987**, *252*, 626−634.
(159) Kettner, C. A.; Shenvi, A. B. *J. Biol. Chem.* **1984**, *259*, 15106−15114.
(160) Matteson, D. S.; Sadhu, K. M.; Lienhard, G. E. *J. Am. Chem. Soc.* **1981**, *103*, 5241−5242.
(161) Bone, R.; Shenvi, A. B.; Kettner, C. A.; Agard, D. A. *Biochemistry* **1987**, *26*, 7609−7614.
(162) Bachovchin, W. W.; Wong, W. Y. L.; Farr-Jones, S.; Shenci, A. B.; Kettner, C. A. *Biochemistry* **1988**, *27*, 7689−7697.
(163) Bone, E.; Frank, D.; Kettner, D.; Agard, D. A. *Biochemistry* **1989**, *28*, 7600−7609.
(164) Bone, R.; Fujishige, A.; Kettner, C. A.; Agard, D. A. *Biochemistry* **1991**, *30*, 10388−10398.
(165) Bone, R.; Silen, J. L.; Agard.D. A. *Nature* **1989**, *339*, 191.
(166) Kettner, D. A.; Bone, R.; Agard, D. A.; Bachovchin, W. W. *Biochemistry* **1988**, *27*, 7682−7688.
(167) Katz, B. A.; Finer-Moore, J.; Mortezaei, R.; Rich, D. H.; Stroud, R. M. *Biochemistry* **1995**, *34*, 8264−8280.
(168) Bernhard, S. A.; Orgel, L. E. *Science* **1959**, *130*, 625−626.
(169) Stroud, R. M.; Kay, L. M.; Dickerson, R. E. *J. Mol. Biol.* **1974**, *83*, 185−208.
(170) Kraut, J. *Annu. Rev. Biochem.* **1977**, *46*, 331−358.
(171) Kossiakoff, A. A.; Spencer.S. A. *Biochemistry* **1981**, *20*, 654−664.
(172) Kossiakoff, A. A.; Spencer, S. A. *Nature* **1980**, *288*, 414−416.
(173) Kossiakoff, A. A.; Spencer, S. A. *Biochemistry* **1981**, *20*, 6462−6474.
(174) Sampson, N. S.; Bartlett, P. A. *Biochemistry* **1991**, *30*, 2255−2263.
(175) Bone, R.; Sampson, N. S.; Bartlett, P. A.; Agard, D. A. *Biochemistry* **1991**, *30*, 2263−2272.
(176) Stubbs, M. T.; Bode, W. *Thromb. Res.* **1993**, *69*, 1.
(177) Markwardt, F. *Methods Enzymol.* **1970**, *19*, 924−932.
(178) Rydel, T. J.; Ravichandran, K. G.; Tulinsky, A.; Bode, W.; Huber, R.; Roitsch, C.; Fenton, J. W. I. *Science* **1990**, *249*, 277−280.
(179) Rydel, T. J.; Tulinski, A.; Bode, W.; Huber, R. *J. Mol. Biol.* **1991**, *221*, 583−601.
(180) Grutter, M. G.; Priestle, M. G.; Rahuel, J.; Grossenbacher, H.; Bode, W.; Hofsteenge, J.; Stone, S. R. *EMBO J.* **1990**, *9*, 2361−2365.
(181) Bode, W.; Turk, D.; Karshikov, A. *Protein Sci.* **1992**, *1*, 426.
(182) Kreutter, K.; Steinmetz, A. C. U.; Liang, T.-C.; Prorok, M.; Abeles, R. H.; Ringe, D. *Biochemistry* **1994**, *33*, 13792−13800.
(183) Angliker, H.; Wikstrom, P.; Rauber, P.; Stone, S.; Shaw, E. *Biochem. J.* **1988**, *256*, 481−486.
(184) Prorok, M.; Albeck, A.; Foxman, B. M.; Abeles, R. H. *Biochemistry* **1994**, *33*, 9784−9790.
(185) Weiner, H.; White, W. N.; Hoare, D. G.; Koshland, D. E., Jr. *J. Am. Chem. Soc.* **1966**, *88*, 3851−3859.
(186) Lim, M. S. L.; Johnston, E. R.; Kettner, C. A. *J. Med. Chem.* **1993**, *36*, 1831.
(187) Bajusz, S.; Szell, E.; Bagly, D.; Barabas, E.; Horvath, G.; Dioszeg, M.; Fittler, Z.; Szabo, G.; Juhasz, A.; Tomori, E.; Szilagyi, G. *J. Med. Chem.* **1990**, *33*, 1729.
(188) Bajusz, S.; Szell, H.; Barabas, E.; Bagdy, D.; Nagy, Z. M. US Patent 4,346,078, 1982.
(189) Wiley, M. R.; Chirgadze, N. Y.; Clawson, D. K.; Craft, T. J.; Gifford-Moore, D. S.; Jones, N. D.; Olkowski, J. L.; Schacht, A. L.; Weir, L. C.; Smith, G. F. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2835−2840.
(190) Weber, P. C.; Lee, S.-L.; Lewandowski, F. A.; Schadt, M. C.; Chang, C.-H.; Kettner, C. A. *Biochemistry* **1995**, *34*, 3750−3757.
(191) Perona, J. J.; Tsu, C. A.; McGrath, M. E.; Craik, C. S.; Fletterick, R. J. *J. Mol. Biol.* **1993**, *230*, 934.
(192) Fevig, J. M.; Abelman, M. M.; Brittelli, D. R.; Kettner, C. A.; Knabb, R., M.; Weber, P. C. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 295−300.
(193) Cacciola, J.; Fevig, J. M.; Alexander, R. S.; Brittelli, D. R.; Kettner, C. A.; Knabb, R. M.; Weber, P. C. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 301−306.
(194) Kettner, C.; Mersinger, L.; Knabb, R. *J. Biol. Chem.* **1990**, *265*, 18289−18297.
(195) Nienaber, V. L.; Mersinger, L. J.; Kettner, C. A. *Biochemistry* **1996**, *35*, 9690−9699.
(196) Burley, S. K.; Petsko, G. A. *Adv. Protein Chem.* **1988**, *39*, 125−189.
(197) Banner, D. W.; Hadvary, P. *J. Biol. Chem* **1991**, *266*, 20085.
(198) Chen, Z.; Li, Y.; Mulichak, A. M.; Lewis, S. D.; Shafer, J. A. *Arch. Biochem. Biophys.* **1995**, *322*, 198.
(199) Kikumoto, R.; Tamao, Y.; Tezuka, T.; Tonomura, S.; Hara, H.; Ninomiya, K.; Hijikata, A.; Okamoto, S. *Biochemistry* **1984**, *23*, 85−90.
(200) Sturzebecher, J.; Markwardt, F.; Voight, B.; Wagner, G.; Walsmann, P. *Thromb. Res.* **1983**, *29*, 635−642.
(201) Suerzebecher, J.; Martin, P. D.; Edwards, B. F. D.; Bode, W. *J. Mol.Biol.* **1992**, *226*, 1085.
(202) Bode, W.; Turk, D.; Stuerzebecher, J. *Eur. J. Biochem.* **1990**, *193*, 175.
(203) Turk, D.; Suerzebecher, J.; Bode, W. *FEBS Lett.* **1991**, *287*, 133.
(204) Matsuzaki, T.; Sasaki, C.; Umeyama, H. *J. Biochem. (Tokyo)* **1988**, *103*, 537−543.
(205) Matsuzaki, T.; Osano, Y. T. *Anal. Sci.* **1989**, *5*, 123.
(206) Brandstetter, H.; Turk, D.; Hoefken, H. W.; Grosse, D.; Stuerzebecher, J.; Martin, P. D.; Edwards, B. F. P.; Bode, W. *J. Mol. Biol.* **1992**, *226*, 1085.
(207) Mack, H.; Pfeiffer, T.; Hornberger, W.; Bohm, H. J.; Hoffken, H. W. *J. Enzyme Inhib.* **1995**, *9*, 73−86.
(208) Iwanoxicz, E. J.; Lau, W. F.; Lin, J.; Roberts, D. G. M.; Seiler, S. M. *J. Med. Chem.* **1994**, *37*, 2122−2124.
(209) Tabernero, L.; Chang, C. Y.; Ohringer, S. L.; Lau, W. F.; Iwanowicz, E. J.; Han, W.-C.; Wang, T. C.; Seiler, S. M.; Roberts, D. G. M.; Sack, J. S. *J. Mol. Biol.* **1995**, *246*, 14−20.
(210) Malley, M. F.; Tabernero, L.; Chang, C. Y.; Ohringer, S. I.; Roberts, G. M.; Das, J.; Sack, J. S. *Protein Sci.* **1996**, *5*, 221−228.
(211) Qui, X.; Padmanabhan, K. P.; Carperos, V. E.; Tulinsky, A.; Kline, T.; Maraganore, J. M.; Fenton, J. W., II. *Biochemistry* **1992**, *31*, 11689.
(212) Krishnan, R.; Tulinsky, A.; Vlasuk, G. P.; Pearson, D.; Vallar, P.; Bergum, P.; Brunck, T. K.; Ripka, W. C. *Protein Sci.* **1996**, *5*, 422−433.
(213) Schultz, R. M.; Varma-Nelson, P.; Ortiz, R.; Kozlowski, K. A.; Orawski, A. T.; Pagast, P.; Frankfater, A. *J. Biol. Chem.* **1989**, *264*, 1497−1507.
(214) Steinmetzer, T.; Konishi, Y. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1677−1682.
(215) Zdanov, A.; Wu, S.; DiMaio, J.; Konishi, Y.; Li, Y.; Wu, X.; Edwards, B. F. P.; Martin, P. D.; Cygler, M. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 252−265.
(216) Bernstein, P. R.; Edwards, P. E.; Williams, J. C. *Prog. Med. Chem.* **1994**, *31*, 59.
(217) Bode, W.; Wei, A. Z.; Huber, R.; Meyer, E.; Travis, J.; Neuman, S. *EMBO J.* **1986**, *5*, 2453−2458.
(218) Navia, M. A.; McKeever, B. M.; Springer, J. P.; Lin, T. Y.; Williams, H. R.; Fluder, E. M.; Dorn, C. P.; Hoogsteen, K. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 7−11.
(219) Meyer, E.; Cole, G.; Radhakrishnan, R.; Epp, O. *Acta Crystallogr., Sect. B* **1988**, *44*, 26.
(220) Edwards, P. E.; Bernstein, P. R. *Med. Res. Rev.* **1994**, *14*, 127−194.
(221) Warner, P.; Green, R. C.; Gomes, B.; Strimper, A. M. *J. Med. Chem.* **1994**, *37*, 3090−3099.
(222) Takahashi, L. H.; Radhakrishan, R.; Rosenfeld, R. E., Jr.; Meyer, E. F., Jr.; Trainer, D. A. *J. Am. Chem. Soc.* **1989**, *111*, 3368−3374.
(223) Edwards, P. D.; Meyer, E. F., Jr.; Vijayalakshmi, J.; Tuthill, P. A.; Andisk, D. A.; Gomes, B.; Strimpler, A. *J. Am. Chem. Soc.* **1992**, *114*, 1854−1863.
(224) Edwards, P. D.; Wolanin, D. J.; Andisik, D. W.; Davis, M. W. *J. Med. Chem.* **1995**, *38*, 76−85.
(225) Edwards, P. D.; Zottola, M. A.; Davis, M.; Williams, J.; Tuthill, P. A. *J. Med. Chem.* **1995**, *38*, 3972−82.
(226) Costanzo, M. J.; Maryanoff, B. E.; Hecker, L. R.; Schott, M. R.; Yabut, S. C.; Zhang, H.-C.; Andrade-Gordon, P.; Kauffman, J. A.; Lewis, J. M.; Krishnan, R.; Tulinsky, A. *J. Med. Chem.* **1996**, *39*, 3039−3043.

(227) Renaud, A.; Lestienne, P.; Hughes, D. L.; Bieth, J. G.; Dimicoli, J.-L. *J. Biol. Chem.* **1983**, *258*, 8312−8316.
(228) Mattos, C.; Giammona, D. A.; Petsko, G. A.; Ringe, D. *Biochemistry* **1995**, *34*, 3193−3203.
(229) Brown, F. J.; Andisik, D. W.; Bernstein, P. R.; Bryant, C. B.; Ceccarelli, C.; Damewood, J. R., Jr.; Edwards, P. D.; Earley, R. A.; Feeney, S.; Green, R. C.; Gomes, B.; Kosmider, B. J.; Krell, R. D.; Shaw, A.; Steelman, G. B.; Thomas, R. M.; Vacek, E. P.; Veale, C. A.; Tuthill, P. A.; Warner, P.; Williams, J. C.; Wolanin, D. J.; Woolson, S. A. *J. Med. Chem.* **1994**, *37*, 1259−1261.
(230) Warner, P.; Green, R. C.; Gomes, B.; Strimpler, A. M. *J. Med. Chem.* **1994**, *37*, 3090−9.
(231) Damewood, J. R., Jr.; Edwards, P. D.; Feeney, S.; Gomes, B. C.; Steelman, G. B.; Tuthill, P. A.; Williams, J. C.; Warner, P.; Woolson, S. A.; Wolanin, D. J.; Veale, C. A. *J. Med. Chem.* **1994**, *37*, 3303−3312.
(232) Bernstein, P. R.; Andisik, D.; Bradley, P. K.; Bryant, C. B.; Ceccarelli, C.; Damewood, J. R., Jr.; Earley, R.; Edwards, P. D.; Feeney, S.; Gomes, B. C.; Kosmider, B. J.; Steelman, G. B.; Thomas, R. M.; Vacek, E. P.; Veale, C. A.; Williams, J. C.; Wolanin, D. J.; Woolson, S. A. *J. Med. Chem.* **1994**, *37*, 3313−3326.
(233) Veale, C. A.; Bernstein, P. R.; Bryant, C.; Ceccarelli, C.; Damewood, J. R., Jr.; Earley, R.; Feeney, S. W.; Gomes, B.; Kosmider, B. J.; Steelman, G. B.; Thomas, R. M.; Vacek, E. P.; Williams, J. C.; Wolanin, D. J.; Woolson, S. *J. Med. Chem.* **1995**, *38*, 98−108.
(234) Bernstein, P. R.; Gomes, B. C.; Kosmider, B. J.; Vacek, E. P.; Williams, J. C. *J. Med. Chem.* **1995**, *38*, 212−215.
(235) Edwards, P. D.; Andisik, D. W.; Strimpler, A. M.; Gomes, B.; Tuthill, P. A. *J. Med. Chem.* **1996**, *39*, 1112−1124.
(236) Navia, M. A.; Springer, J. P.; Lin, T.-Y.; Williams, H. R.; Firestone, R. A.; Pisano, J. M.; Doherty, J. B.; Finke, P. E.; Hoogstein, K. *Nature* **1987**, *327*, 79−82.
(237) Green, B. G.; Chabin, R.; Mills, S.; Underwood, D. J.; Shah, S. K.; Kuo, D.; Gale, P.; Maycock, A. L.; Liesch, J.; Burgey, C. S.; Doherty, J. B.; Dorn, C. P.; Finke, P. E.; Hagmann, W. K.; Hale, J. J.; MacCoss, M.; Westler, W. M.; Knight, W. B. *Biochemistry* **1995**, *34*, 14331−14343.
(238) Radhakrishnan, R.; Presta, L. G.; Meyer, E. F.; Wildonger, R. *J. Mol. Biol.* **1987**, *198*, 417−424.
(239) Ghuysen, J. M. *Annu. Rev. Microbiol.* **1991**, *154*, 37−67.
(240) Jelsch, C.; Mourey, L.; Masson, J. M.; Samma, J. P. *Proteins: Struct. Funct. Genet.* **1993**, *16*, 364.
(241) Hardy, L. W.; Kirsch, J. F. *Biochemistry* **1984**, *23*, 1275.
(242) Adachi, H.; Ohta, T.; Matsuzawa, H. *J. Biol. Chem.* **1991**, *266*, 3186.
(243) Escobar, W. A.; Tan, A. K.; Fink, A. L. *Biochemistry* **1991**, *30*, 10783.
(244) Delaire, M.; Lenfant, F.; Labia, R.; Masson, J. M. *Protein Eng.* **1991**, *4*, 805.
(245) Rizwi, I.; Tan, A. K.; Fink, A. L.; Virden, R. *Biochem. J.* **1989**, *258*, 205−209.
(246) Chen, C. C.; Herzberg *J. Mol. Biol.* **1992**, *224*, 1103−1113.
(247) Martin, R.; Jones, J. B. *Tetrahedron Lett.* **1995**, *36*, 8399−8402.
(248) Strynadka, N. C. J.; Martin, R.; Jensen, S. E.; Gold, M.; Jones, J. B. *Nat. Struct. Biol.* **1996**, *3*, 688−695.
(249) Miyashita, K.; Massova, I.; Taibi, P.; Mobashery, S. *J. Am. Chem. Soc.* **1995**, *117*, 11055−11059.
(250) Maveyraud, L.; Massova, I.; Birck, C.; Miyashita, K.; Samama, J.-P.; Mobashery, S. *J. Am. Chem. Soc.* **1996**, *118*, 7435−7440.
(251) Goldberg, A. L. *Science* **1995**, *268*, 522−523.
(252) Spaltenstein, A.; Leban, J. J.; Huang, J. J.; Reinhardt, K. R.; Viveros, O. H.; Sigafoos, J.; Crouch, R. *Tetrahedron Lett.* **1996**, *37*, 343−346.
(253) Palombella, V. J.; Rando, O. J.; Goldberg, A. L.; Maniatis, T. *Cell* **1994**, *78*, 773.
(254) Treier, M.; Stotzewski, L. M.; Bohmann, D. *Cell* **1994**, *78*, 878.
(255) Fenteany, G.; Standaert, R. F.; Lane, W. S.; Choi, S.; Corey, E. J.; Schreiber, S. L. *Science* **1995**, *268*, 726−731.
(256) Lowe, J.; Stock, D.; Jap, B.; Zwicki, P.; Baumeister, W.; Huber, R. *Science* **1995**, *268*, 533−539.
(257) Adams, J.; Stein, R. *Annu. Rep. Med. CHem.* **1996**, *31*, 279−288.
(258) Shaw, E. *Adv. Enymol. Relat. Area Mol. Biol.* **1990**, *63*, 271−347.
(259) Rich, D. H. In *Proteinase Inhibitors*; Barrett, A. J., Salveson, G., Eds.; Elsevier: Amsterdam, 1986; Vol. 12, pp 153−178.
(260) Shaw, E. *J. Protein Chem.* **1984**, *3*, 109−119.
(261) Otto, H.-H.; Schirmeister, T. *Chem. Rev.* **1997**, *97*, 133−171.
(262) Lewis, S. D.; Johnson, F. A.; Shafer, J. A. *Biochemistry* **1976**, *15*, 5009.
(263) Polgar, L.; Halasz, P. *Biochem. J.* **1982**, *207*, 1−10.
(264) O'Leary, M. H.; Urberg, M.; Young, A. P. *Biochemistry* **1974**, *13*, 2077−2081.
(265) Wagner, B. M.; Smith, R. A.; Coles, P. J.; Cobb, L. J.; Ernest, M. J.; Krantz, A. *J. Med. Chem.* **1994**, *37*, 1833−1840.
(266) Kamphius, I. G.; Drenth, J.; Baker, E. N. *J. Mol. Biol.* **1985**, *182*, 317−329.
(267) Takio, K.; Towatari, T.; Katanuma, N.; Teller, D. C.; Titani, K. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 3666−3670.
(268) Schroder, E.; Phillips, C.; Garman, E.; Harlos, K.; Crawford, C. *FEBS Lett.* **1993**, *315*, 38−42.
(269) Gamcsik, M. P.; Malthouse, J. P. G.; Primrose, W. U.; Mackenzie, N. E.; Boyd, A. S. F.; Russell, R. A.; Scott, A. I. *J. Am. Chem. Soc.* **1983**, *105*, 6324−6325.
(270) Smith, R. A.; Copp, L. J.; Donnelly, S. L.; Spencer, R. W.; Krantz, A. *Biochemistry* **1988**, *27*, 6568−6573.
(271) Smith, R. A.; Copp, L. J.; Coles, P. J.; Pauls, H. W.; Robinson, V. J.; Spencer, R. W.; Heard, S. B.; Krantz, A. *J. Am. Chem. Soc.* **1988**, *110*, 4429−4431.
(272) Lewis, C. A.; Wolfenden, R. *Biochemistry* **1977**, *16*, 4890−4895.
(273) Westerick, J. O.; Wolfenden, R. *J. Biol. Chem.* **1972**, *247*, 8195−8197.
(274) Thompson, S. A.; Andrews, P. R.; Hanzlik, R. P. *J. Med. Chem.* **1986**, *29*, 104−111.
(275) Moon, J. B.; Coleman, R. S.; Hanzlik, R. P. *J. Am. Chem. Soc.* **1986**, *108*, 1350−1351.
(276) Dufour, E.; Storer, A. C.; Menard, R. *Biochemistry* **1995**, *34*, 9136−9143.
(277) Hu, L. Y.; Abeles, R. H. *Arch. Biochem. Bophys.* **1990**, *281*, 271−274.
(278) Saido, T. C.; Sorimachi, H.; Suzuki, K. *FASEB J.* **1994**, *8*, 814−822.
(279) Berti, P. J.; Storer, A. C. *J. Mol. Biol.* **1995**, *246*, 273−283.
(280) Li, Z.; Patil, G. S.; Golubski, Z. E.; Hori, H.; Tehrani, K.; Foreman, J. E.; Eveleth, D. D.; Bartus, R. T.; Powers, J. C. *J. Med. Chem.* **1993**, *36*, 3472−3480.
(281) Drenth, J.; Kalk, K. H.; Swen, H. M. *Biochemistry* **1976**, *15*, 3731−3738.
(282) O'Hara, B. P.; Hemmings, A., M.; Buttle, D. J.; Pearl, L. H. *Biochemistry* **1995**, *34*, 13190−13195.
(283) Rasnick, D. *Anal. Biochem.* **1985**, *149*, 461−465.
(284) Shaw, E. *J. Biol. Chem.* **1988**, *263*, 2768−2772.
(285) Varughese, K. I.; Ahmed, F. R.; Carey, P. R.; Hasnain, S.; Huber, C. P.; Storer, A. C. *Biochemistry* **1989**, *28*, 1330−1332.
(286) Varughese, K. I.; Su, Y.; Cromwell, D.; Hasnain, S.; Xuong, N.-h. *Biochemistry* **1992**, *31*, 5172−5176.
(287) Yamamoto, D.; Matsumoto, K.; Ohishi, H.; Ishida, T.; Inoue, M.; Kitamura, K.; Mizuno, H. *J. Biol. Chem.* **1991**, *266*, 14771.
(288) Kim, M.-J.; Yamamoto, D.; Matsumoto, K.; Inoue, M.; Ishida, T.; Mizuno.H.; Sumiya, S.; Kitamura, K. *Biochem. J.* **1992**, *287*, 797.
(289) Turk, D.; Podobnik, M.; Popovic, T.; Katunuma, N.; Bode, W.; Huber, R.; Turk, V. *Biochemistry* **1995**, *34*, 4791−4797.
(290) Meara, J. P.; Rich, D. H. *J. Med. Chem.* **1996**, *39*, 3357−3366.
(291) Palmer, J. T.; Rasnick, D.; Klaus, J. L.; D., B. *J. Med. Chem.* **1995**, *38*, 3193−3196.
(292) Thornberry, N. A.; Miller, D. K.; Nicholson, D. W. *Perspect. Drug Discovery Des.* **1994**, *2*, 389−399.
(293) Walker, N. P. C.; Talanian, R. V.; Brady, K. D.; Dang, L. C.; Bump, N. J.; Ferenz, C. R.; Franklin, S.; Ghayur, T.; Hackett, M. C.; Hammill, L. D.; Herzog, L.; Hugunin, M.; Houy, W.; Mankovich, J. A.; McGuiness, L.; Orlewicz, E.; Paskind, M.; Pratt, C. A.; Reis, P.; Summani, A.; Terranova, M.; Welch, J. P.; Xiong, L.; Moller, A.; Tracey, D. E.; Kamen, R.; Wong, W. W. *Cell* **1994**, *78*, 343−352.
(294) Wilson, K. P.; Black, J.-A., F.; Thomson, J. A.; Kim, E. E.; Griffith, J. P.; Navia, M. A.; Murcko, M. A.; Chambers, S. P.; Aldape, R. A.; Raybuck, S. A.; Livingston, D. J. *Nature* **1994**, *370*, 270−275.
(295) Thornberry, N. A.; Bull, H. G.; Calaycay, J. R.; Chapman, K. T.; Howard, A. D.; Kostura, M. J.; Miller, D. K.; Molineaux, S. M.; Weidner, J. R.; Aunins, J.; Elliston, K. O.; Ayala, J. M.; Casano, F. J.; Chin, J.; Ding, G. J.-F.; Egger, L. A.; Gaffney, E. P.; Limjuco, G.; Palyha, O. C.; Raju, S. M.; Rolando, A. M.; Salley, J. P.; Yamin, T.-T.; Lee, T. D.; Shively, J. E.; MacCoss, M.; Mumford, R. A.; Schmidt, J. A.; Tocci, M. J. *Nature* **1992**, *356*, 768.
(296) Chapman, K. T. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 613.
(297) Graybill, T. L.; Dolle, R. E.; Helaszek, C. T.; Miller, R. E.; Ator, M. A. *Int. J. Pept. Protein Res.* **1994**, *44*, 173.
(298) Mullican, M. D.; Lauffer, D. J.; Gillespie, R. J.; Matharu, S. S.; Kay, D.; Porritt, G. M.; Evans, P. L.; Golec, J. M.; Murcko, M. A.; Luong, Y.-P.; Raybuck, S. A.; Livingston, D. J. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2359.
(299) Thornberry, N. A.; Molineaux, S. M. *Protein Sci.* **1995**, *4*, 3−12.
(300) Robinson, R. P.; Donahue, K. M. *J. Org. Chem.* **1992**, *57*, 7309−7314.
(301) Mjalli, A. M. M.; Chapman, K. T.; MacCoss, M.; Thornberry, N. A. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 2689−2692.
(302) Mjalli, A. M. M.; Chapman, K. T.; MacCoss, M. A. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 2693−2698.
(303) Mjalli, A. M. M.; Chapman, K. T.; MacCoss, M.; Thornberry, N. A.; Peterson, E. P. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1965.
(304) Hellen, C. U. T.; Krausslich, H.-G.; Wimmer, E. *Biochemistry* **1989**, *28*, 9881−9889.
(305) Malcom, B. A. *Protein Sci.* **1995**, *4*, 1439−1445.
(306) Bazan, J. F.; Fletterick, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7872−7876.
(307) Gorbalenya, A. E.; Donchecko, A. P.; Blinov, V. M.; Koomin, E. V. *FEBS Lett.* **1989**, *243*, 103−114.

(308) Matthews, D. A.; Smith, W. W.; Ferre, R. A.; Condon, B.; Budahazi, G.; Sisson, W.; Villafranca, J. E.; Janson, C. A.; McElroy, H. E.; Gribskov, C. L.; Worland, S. *Cell* **1994**, *77*, 1−20.
(309) Allaire, M.; Chernaia, M. M.; Malcolm, B. A.; James, M. N. G. *Nature* **1994**, *369*, 72−76.
(310) Higaki, J. N.; Evnin, L. B.; Craik, C. S. *Biochemistry* **1989**, *28*, 9256−9263.
(311) McGrath, M. E.; Wilke, M. E.; Higaki, J. N.; Craig, C. S.; Fletterick, R. J. *Biochemistry* **1989**, *28*, 9264−9270.
(312) Nakatsuka, T.; Sasaki, T.; Kaiser, E. T. *J. Am. Chem. Soc.* **1987**, *109*, 3808−3810.
(313) Wu, Z.-P.; Hilvert, D. *J. Am. Chem. Soc.* **1989**, *111*, 4513−4514.
(314) Syed, R.; Wu, Z.-P.; Hogle, J. M.; Hilvert, D. *Biochemistry* **1993**, *32*, 6157.
(315) Jackson, D. Y.; Burnier, J.; Quan, C.; Stanley, M.; Tom, J.; Wells, J. A. *Science* **1994**, *266*, 243−247.
(316) Abrahamsen, L.; Tom, J.; Burnier, K. A.; Kossiakoff, A.; Wells, J. A. *Biochemistry* **1991**, *30*, 4151−4159.
(317) Lawson, M. A.; Semler, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9919−9923.
(318) Kaldor, S. W.; Hammond, M.; Dressman, B. A.; Labus, J. M.; Chadwell, F. W.; Kline, A. D.; Heinz, B. A. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2021−2106.
(319) Kline, A. D. Manuscript in preparation, cited in ref 318.
(320) Malcolm, B. A.; Lowe, C.; Shechosky, S.; McKay, R. T.; Yang, C. C.; Shah, V. J.; Simon, R. J.; Vederas, J. C.; Santi, D. V. *Biochemistry* **1995**, *34*, 8172−8179.
(321) Webber, S. E.; Tikhe, J.; Worland, S. T.; Fuhrman, S. A.; Hendrickson, T. F.; Matthews, D. A.; Love, R. A.; Patick, A. K.; Meador, J. W.; Ferre, R. A.; Brown, E. L.; DeLise, D. M.; Ford, C. E.; Binford, S. L. *J. Med. Chem.* **1996**, *39*, 5072−5082.
(322) Menard, R.; Plouffe, C.; Laflamme, P.; Vernet, T.; Tessier, D. C.; Thomas, D. Y.; Storer, A. C. *Biochemistry* **1995**, *34*, 464−471.
(323) Asboth, B.; Stokum, E.; Khan, I. U.; Polgar, L. *Biochemistry* **1985**, *24*, 606−9.
(324) Robinson, V. J.; Pauls, H. W.; Coles, P. J.; Smith, R. A.; Krantz, A. *Bioorg. Chem.* **1992**, *20*, 42−54.
(325) Krantz, A.; Copp, L. J.; Coles, P. J.; Smith, R. A.; Heard, S. B. *Biochemistry* **1991**, *30*, 4678−4687.
(326) Dolle, R. E.; Hoyer, D.; Prasad, C. V. C.; Schmidt, S. J.; Helaszek, C. T.; Miller, R. E.; Ator, M. A. *J. Med. Chem.* **1994**, *37*, 563−564.
(327) Dolle, R. E.; Singh, J.; Rinker, J.; Hoyer, D.; Prasad, C. V. C.; Graybill, T. L.; Salvino, J. M.; Helaszek, C. T.; Miller, R. E.; Ator, M. A. *J. Med. Chem.* **1994**, *37*, 3863−3866.
(328) Dolle, R. E.; Singh, J.; Whipple, D.; Osifo, I. K.; Speier, G.; Graybill, T. L.; Gregory, J. S.; Harris, A. L.; Helaszek, C. T.; Miller, R. E.; Ator, M. A. *J. Med. Chem.* **1995**, *38*, 220−222.
(329) Stöcker, W.; Bode, W. *Curr. Opin. Struct. Biol.* **1995**, *5*, 383−390.
(330) Stöcker, W.; Grams, F.; Baumann, U.; Reinemer, P.; Gomis-Rüth, F.-X.; McKay, D. B.; Bode, W. *Protein Sci.* **1995**, *4*, 823−840.
(331) Bode, W. *Structure* **1995**, *3*, 527.
(332) Blundell, T. L. *Nature Struct. Biol.* **1994**, *1*, 73−75.
(333) Roques, B. P. *Biochem. Soc. Trans.* **1993**, *21*, 678−685.
(334) Lipscomb, W. N.; Sträter, N. *Chem. Rev.* **1996**, *96*, 2375−2433.
(335) Christianson, D. W.; Lipscomb, W. N. *Acc. Chem. Res.* **1989**, *22*, 62−69.
(336) Feinberg, H.; Greenblatt, H. M.; Behar, V.; Gilon, C.; Cohen, S.; Bino, A.; Shoham, G. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1995**, *D51*, 428−49.
(337) Matthews, B. W. *Acc. Chem. Res.* **1988**, *21*, 333−340.
(338) Wyvratt, M. J.; Patchett, A. A. In *Medicinal Research Reviews*; John Wiley & Sons, Inc.: New York, 1985; Vol. 5, pp 483−531.
(339) Waller, C. L.; Marshall, G. R. *J. Med. Chem.* **1993**, *36*, 2390−2403.
(340) Keller, P. M.; Lee, C.-P.; Fenwick, A. E.; Atkinson, S. T.; Elliott, J. D.; DeWolf, W. E., Jr. *Biochem. Biophys. Res. Commun.* **1996**, *223*, 372−378.
(341) Bihovsky, R.; Levinson, B. L.; Loewi, R. C.; Erhardt, P. W.; Polokoff, M. A. *J. Med. Chem.* **1995**, *38*, 2119−29.
(342) Lago, M. A.; Luengo, J. I.; Peishoff, C. E.; Elliot, J. D. *Annu. Rep. Med. Chem.* **1996**, *31*, 81−90.
(343) Woessner, J. F., Jr. *FASEB J.* **1991**, *5*, 2145−2154.
(344) Hagmann, W. K.; Lark, M. W.; Becker, J. W. *Annu. Rep. Med. Chem* **1996**, *31*, 231−240.
(345) Stams, T.; John, C. S.; Smith, D. L.; Wahl, R. C.; Ho, T. F.; Qoronfleh, M. W.; Banks, T. M.; Rubin, B. *Nat. Struct. Biol.* **1994**, *1*, 119−23.
(346) Spurlino, J. C.; Smallwood, A. M.; Carlton, D. D.; Banks, T. M.; Vavra, K. J.; Johnson, J. S.; Cook, E. R.; Falvo, J.; Wahl, R. C.; et al. *Proteins: Struct., Funct., Genet.* **1994**, *19*, 98−109.
(347) Chen, J. J.; Zhang, Y.; Hammond, S.; Dewdney, N.; Ho, T.; Lin, X.; Browner, M. F.; Castelhano, A. L. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1601−1606.
(348) Browner, M. F.; Smith, W. W.; Castelhano, A. L. *Biochemistry* **1995**, *34*, 6602−10.
(349) Lovejoy, B.; Cleasby; A.; Hassell, A. M.; Longley, K.; Luther, M. A.; Weigl, D.; McGeehan, G.; McElroy, A. B.; Drewry, D.; Lambert, M. H.; Jordan, S. R. *Science* **1994**, *263*, 375−377.
(350) Nagase, H.; Fields, G. B. *Biopolymers* **1996**, *40*, 399−416.
(351) Phillips, M. A.; Fletterick, R.; Rutter, W. J. *J. Biol. Chem.* **1990**, *265*, 20692−20698.
(352) Lovejoy, B.; Hassell, A. M.; Luther, M. A.; Weigl, D.; Jordan, S. R. *Biochemistry* **1994**, *33*, 8207−8217.
(353) Holland, D. R.; Hausrath, A. C.; Juers, D.; Matthews, B. W. *Protein Sci.* **1995**, *4*, 1955.
(354) Appelt, K.; Almassy, R. Unpublished results.
(355) Beckett, R. P.; Davidson, A. H.; Drummond, A. H.; Huxley, P.; Whittaker, M. *Drug Discovery Today* **1996**, *1*, 16−26.
(356) Porter, J. R.; Millican, T. A.; Morphy, J. R. *Expert Opin. Ther. Pat.* **1995**, *5*, 1287−96.
(357) Morphy, J. R.; Millican, T. A.; Porter, J. R. *Curr. Med. Chem.* **1995**, *2*, 743−62.
(358) Browner, M. F. *Perspect. Drug Discovery Des.* **1995**, *2*, 343−51.
(359) Schwartz, M. A.; Van Wart, H. E. *Prog. Med. Chem.* **1992**, *29*, 271−334.
(360) Castelhano, A. L.; Billedeau, R.; Dewdney, N.; Donnelly, S.; Horne, S.; Kurz, L. J.; Liak, T. J.; Martin, R.; Uppington, R.; et al. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1415−20.
(361) Morphy, J. R.; Beeley, N. R. A.; Boyce, B. A.; Leonard, J.; Mason, B.; Millican, A.; Millar, K.; O'Connell, J. P.; Porter, J. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2747−2752.
(362) Caldwell, C. G.; Sahoo, S. P.; Polo, S. A.; Eversole, R. R.; Lanza, T. J.; Mills, S. G.; Niedzwiecki, L. M.; Izquierdo-Martin, M.; Chang, B. C.; et al. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 323−8.
(363) Goulet, J. L.; Kinneary, J. F.; Durrette, P. L.; Stein, R. L.; Harrison, R. K.; Izquierdo-Martin, M.; Kuo, D. W.; Lin, T.-Y.; Hagmann, W. K. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1221−1224.
(364) Porter, J. R.; Beeley, N. R. A.; Boyce, B. A.; Mason, B.; Millican, A.; Millar, K.; Leonard, J.; Morphy, J. R.; O'Connell, J. P. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2741−2746.
(365) Tomczuk, B. E.; Gowravaram, M. R.; Johnson, J. S.; Delecki, D.; Cook, E. R.; Ghose, A. K.; Mathiowetz, A. M.; Spurlino, J. C.; Rubin, B.; Smith, D. L.; Pulvino, T.; Wahl, R. C. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 343−348.
(366) Pease, J.; Bender, S. L. Unpublished results.
(367) Holmes, M. A.; Matthews, B. W. *Biochemistry* **1981**, *20*, 6912−6920.
(368) Bender, S. L.; Deal, J. G.; Zook, S. E.; Appelt, K.; Almassy, R. J.; Abreo, M. A.; Agree, C. S.; Melnick, M. J.; Mitchell, L.; Margosiak, S. A.; Dagostino, E. J.; Wickersham, J. Manuscript in preparation.
(369) Castelhano, A. L.; Almassy, R. J.; Wickersham, J.; Appelt, K. Unpublished results.
(370) Browner, M. F. Unpublished results.
(371) Kim, H.; Lipscomb, W. N. *Biochemistry* **1990**, *29*, 5546−55.
(372) Kim, H.; Lipscomb, W. N. *Biochemistry* **1991**, *30*, 8171−80.
(373) Cappalonga, A. M.; Alexander, R. S.; Christianson, D. W. *J. Biol. Chem.* **1992**, *267*, 19192−19197.
(374) Grams, F.; Reinemer, P.; Powers, J. C.; Kleine, T.; Pieper, M.; Tschesche, H.; Huber, R.; Bode, W. *Eur. J. Biochem.* **1995**, *228*, 830−41.
(375) Becker, J. W.; Marcy, A. I.; Rokosz, L. L.; Axel, M. G.; Burbaum, J. J.; Fitzgerald, P. M. D.; Cameron, P. M.; Esser, C. K.; Hagmann, W. K.; et al. *Protein Sci.* **1995**, *4*, 1966−76.
(376) Gooley, P. R.; O'Connell, J. F.; Marcy, A. I.; Cuca, G. C.; Axel, m. G.; Caldwell, C. G.; Hagmann, W. K.; Becker, J. W. *J. Biomol. NMR* **1996**, *7*, 8−28.
(377) Abreo, M. A.; Bender, S. L.; Appelt, K.; Agree, C. S.; Zook, S. E.; Almassy, R.; Margosiak, S. A.; Dagostino, E. J. Manuscript in preparation.
(378) Wahl, R. C.; Pulvino, T. A.; Mathiowetz, A. M.; Ghose, A. K.; Johnson, J. S.; Delecki, D.; Cook, E. R.; Gainor, J. A.; Gowravaram, M. R.; Tomczuk, B. E. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 349−352.
(379) Netzel-Arnett, S.; Fields, G.; Birkedal-Hansen, H.; Van Wart, H. *J. Biol. Chem.* **1991**, *266*, 6747−6755.
(380) Sang, Q. A.; Douglas, D. A. *J. Protein Sci.* **1996**, *15*, 137−160.
(381) Almassy, R. J. Personal communication.
(382) Dougherty, D. A. *Science* **1996**, *271*, 163−168.
(383) Johnson, W. H.; Roberts, N. A.; Borkakoti, N. *J. Enzyme Inhib.* **1987**, *2*, 1−22.
(384) Borkakoti, N.; Winkler, F. K.; Williams, D. H.; D'Arcy, A.; Broadhurst, M. J.; Brown, P. A.; Johnson, W. H.; Murray, E. J. *Nat. Struct. Biol.* **1994**, *1*, 106−10.
(385) Dhanaraj, V.; Ye, Q. Z.; Johnson, L. L.; Hupe, D. J.; Ortwine, D. F.; Dunbar, J. B., Jr.; Rubin, J. R.; Pavlovsky, A.; Humblet, C.; Blundell, T. L. *Structure* **1996**, *4* (4), 375−386.
(386) Van Doren, S. R.; Kurochkin, A. V.; Hu, W.; Ye, Q.-Z.; Johnson, L. L.; Hupe, D. J.; Zuiderweg, E. R. P. *Protein Sci.* **1995**, *4*, 2487−98.
(387) Johnson, W. H. *Drug News Perspect.* **1990**, *3*, 453.
(388) Margosiak, S. A.; Dagostino, E. J.; Bender, S. L. Unpublished results.
(389) Grams, F.; Crimmin, M.; Hinnes, L.; Huxley, P.; Pieper, M.; Tschesche, H.; Bode, W. *Biochemistry* **1995**, *34*, 14012−20.
(390) Botos, I.; Scapozza, L.; Zhang, D.; Liotta, L. A.; Meyer, E. F. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 2749−54.

(391) Chapman, K. T.; Durette, P. L.; Caldwell, C. G.; Sperow, K. M.; Niedzwiecki, L. M.; Harrison, R. K.; Saphos, C.; Christen, A. J.; Olszewski, J. M.; et al. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 803−6.

(392) Bird, J.; De Mello, R. C.; Harper, G. P.; Hunter, D. J.; Karran, E. H.; Markwell, R. E.; Miles-Williams, A. J.; Rahman, S. S.; Ward, R. W. *J. Med. Chem.* **1994**, *37*, 158−169.

(393) Xue, C.-B.; He, X.; Roderick, J.; DeGrado, W. F.; Decicco, C.; Copeland, R. A. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 379−84.

(394) Brown, F. K.; Brown, P. J.; Bickett, D. M.; Chambers, C. L.; Davies, H. G.; Deaton, D. N.; Drewry, D.; Foley, M.; McElroy, A. B.; Gregson, M.; McGeehan, G. M.; Myers, P. L.; Norton, D.; Salovich, J. M.; Schoenen, F. J.; Ward, P. *J. Med. Chem.* **1994**, *37*, 674−688.

(395) Esser, C. K.; Bugianesi, R. L.; Caldwell, C. G.; Chapman, K. T.; Durette, P. L.; Girotra, N. N.; Kopka, I. E.; Lanza, T. J.; Levorse, D. A.; MacCoss, M.; Owens, K. A.; Ponpipom, M. M.; Simeone, J. P.; Harrison, R. K.; Niedzwiecki, L.; Becker, J. W.; Marcy, A. I.; Axel, M. G.; Christen, A. J.; McDonnell, J.; Moore, V. L.; Olszewski, J. M.; Saphos, C.; Visco, D. M.; Shen, F.; Colletti, A.; Krieter, P. A.; Hagmann, W. K. *J. Med. Chem.* **1997**, *40*, 1026−1040.

(396) Almassy, R. J.; Horne, S.; Wickersham, J.; Appelt, K. Unpublished results.

(397) Gooley, P. R.; O'Connell, J. F.; Marcy, A. I.; Cuca, G. C.; Salowe, S. P.; Bush, B. L.; Hermes, J. D.; Esser, C. K.; Hagmann, W. K.; Springer, J. P.; Johnson, B. A. *Nature Struct. Biol.* **1994**, *1*, 111−118.

(398) Esser, C. K. personal communication.

(399) MacPherson, L. J.; Parker, D. T. EP 606046, 1994.

(400) Miller, A.; Whittaker, M.; Beckett, R. P. WO 9535276, 1995.

(401) Miller, A.; Whittaker, M.; Beckett, R. P. WO 9535275, 1995.

(402) Parker, D. T.; MacPherson, L. J.; Goldstein, R.; Justice, M. R.; Zhu, L. J.; Choquette, D.; Li, H. J.; Capparelli, M.; Whaley, L. W.; Boehm, C.; O'Byrne, E. M.; Goldberg, R. L.; Ganu, V. S. Poster presented at Inflammation Research Assocation Conference, September 25−29, 1994.

(403) Dreyfuss, M.; Harri, E.; Hofmann, H.; Kobel, H.; Pache, W.; Tscherter, H. *Eur. J. Appl. Microbiol.* **1976**, *3*, 25.

(404) Tanaka, H.; Kuroda, A.; Marusawa, H.; Hatanaka, H.; Kino, T.; Hoto, T.; Hashimoto, M.; Taga, T. *J. Am. Chem. Soc.* **1987**, *109*, 5031−5033.

(405) Sehgal, S. N.; Baker, H.; Vezina, C. *J. Antibiot.* **1975**, *28*, 721.

(406) Vezina, C.; Kudelski, A.; Sehgal, S. N. *J. Antibiot.* **1975**, *28*, 727.

(407) Handschumacher, R. E.; Harding, M. W.; Rice, J.; Drugge, R. J.; Speicher, D. W. *Science* **1984**, *226*, 544−547.

(408) Takahashi, N.; Hayano, T.; Suzuki, M. *Nature* **1989**, *337*, 473−475.

(409) Fischer, G.; Wittmann-Liebold, B.; Lang, K.; Kiefhaber, T.; Schmid, F. X. *Nature* **1989**, *337*, 476−478.

(410) Siekierka, J. J.; Hung, S. H.; Poe, M.; Lin, C. S.; Sigal, N. H. *Nature* **1989**, *341*, 755−757.

(411) Harding, M. W.; Galat, A.; Uehling, D. E.; Schreiber, S. L. *Nature* **1989**, *341*, 758−760.

(412) Wiederrecht, G.; Etzkorn, F. *Perspect. Drug Discovery Des.* **1994**, *2*, 57−84.

(413) Fischer, G. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1425−1436.

(414) Sigal, N. H.; Dumont, F.; Durette, P.; Siekierka, J. J.; Peterson, L.; Rich, D. H.; Dunlap, B. E.; Staruch, M. J.; Melino, M. R.; Koprac, S. L.; Williams, D.; Witzel, B.; Pisano, J. M. *J. Exp. Med.* **1991**, *173*, 619−628.

(415) Schreiber, S. L. *Science* **1991**, *251*, 283−287.

(416) Liu, J.; Farmer, J. D., Jr.; Lane, W. S.; Friedman, J.; Weissman, I.; Schreiber, S. L. *Cell* **1991**, *66*, 807−815.

(417) Aldape, R. A.; Futer, O.; DeCenzo, M. T.; Jarrett, B. P.; Murcko, M. A.; Livingston, D. J. *J. Biol. Chem.* **1992**, *267*, 16029−16032.

(418) Yang, D.; Rosen, M. K.; Schreiber, S. L. *J. Am. Chem. Soc.* **1993**, *115*, 819−820.

(419) Rosen, M. K.; Yang, D.; Martin, P. K.; Schreiber, S. L. *J. Am. Chem. Soc.* **1993**, *115*, 821−822.

(420) Brown, E. J.; Albers, M. W.; Shin, T. B.; Ichikawa, K.; Keith, C. T.; Lane, W. S.; Schreiber, S. L. *Nature* **1994**, *369*, 756−758.

(421) Liu, J.; Albers, M. W.; Wandless, T. J.; Luan, S.; Alberg, D. G.; Belshaw, P. J.; Cohen, P.; MacKintosh, C.; Klee, C. B.; Schreiber, S. L. *Biochemistry* **1992**, *31*, 3896−3901.

(422) Austin, D. J.; Crabtree, G. R.; Schreiber, S. L. *Chem. Biol.* **1994**, *1*, 131−136.

(423) Van Duyne, G. D.; Standaert, R. F.; Schreiber, S. L.; Clardy, J. *J. Am. Chem. Soc.* **1991**, *113*, 7433−7434.

(424) Van Duyne, G. D.; Standaert, R. F.; Karplus, P. A.; Schreiber, S. L.; Clardy, J. *J. Mol. Biol.* **1993**, *229*, 105−124.

(425) Van Duyne, G. D.; Standaert, R. F.; Karplus, P. A.; Schreiber, S. L.; Clardy, J. *Science* **1991**, *252*, 839−842.

(426) Thomas, K. A.; Smith, G. M.; Thomas, T. B.; Feldman, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4843.

(427) Rosen, M. K.; Standaert, R. F.; Galat, A.; Nakatsuka, M.; Schreiber, S. L. *Science* **1990**, *248*, 863−866.

(428) Albers, M. W.; Walsh, C. T.; Schreiber, S. L. *J. Org. Chem.* **1990**, *55*, 4984−4986.

(429) Harrison, R. K.; Stein, R. L. *Biochemistry* **1990**, *29*, 1684−1689.

(430) Liu, J.; Albers, M. W.; Chen, C.-M.; Schreiber, S. L.; Walsh, C. T. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 2304−2308.

(431) Holt, D. A.; Konialian-Beck, A. L.; Oh, H.-J.; Yen, H.-K.; Rozamus, L. W.; Krog, A. J.; Erhard, K. F.; Ortiz, E.; Levy, M. A.; Brandt, M.; Bossard, M. J.; Luengo, J. I. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 315−320.

(432) Yamashita, D. S.; Oh, H.-J.; Yen, H.-K.; Bossard, M. J.; Brandt, M.; Levy, M. A.; Newman-Tarr, T.; Badger, A.; Luengo, J. I.; Holt, D. A. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 325−328.

(433) Luengo, J. I.; Konialian-Beck, A.; Levy, M. A.; Brandt, M.; Eggleston, D. S.; Holt, D. A. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 21−24.

(434) Holt, D. A.; Luengo, J. I.; Yamashita, D. S.; Oh, H. J.; Konialian, A. L.; Yen, H. K.; Rozamus, L. W.; Brandt, M.; Bossard, M. J.; Levy, M. A.; Eggleston, D. S.; Liang, J.; Schultz, L. W.; Stout, T. J.; Clardy, J. C. *J. Am. Chem. Soc.* **1993**, *115*, 9925−9938.

(435) Armistead, D. M.; Badia, M. C.; Deininger, D. D.; Duffy, J. P.; Saunders, J. O.; Tung, R. D.; Thomson, J. A.; DeCenzo, M. T.; Futer, O.; Livingston, D. J.; Murcko, M. A.; Yamashita, M. M.; Navia, M. A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1995**, *D51*, 522−528.

(436) Katoh, S.; Parge, H.; Kissinger, C.; Lewis, C. Unpublished results.

(437) Wang, G. T.; Lane, B.; Fesik, S. W.; Petros, A.; Luly, J.; Krafft, G. A. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1161−1166.

(438) Tatlock, J.; Kalish, V.; Parge, H.; Kissinger, C.; Lewis, C. Unpublished results.

(439) Tatlock, J. H.; Kalish, V. J.; Parge, H. E.; Knighton, D. R.; Showalter, R. E.; Lewis, C. T.; French, J. V.; Villafranca, J. E. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2489−2494.

(440) Dragovich, P. S.; Barker, J. E.; French, J.; Imbacuan, M.; Kalish, V. J.; Kissinger, C. R.; Knighton, D. R.; Lewis, C. T.; Moomaw, E. W.; Parge, H. E.; Pelletier, L. A. K.; Prins, T. J.; Showalter, R. E.; Tatlock, J. H.; Tucker, K. D.; Villafranca, J. E. *J. Med. Chem.* **1996**, *39*, 1872−1884.

(441) Andres, C. J.; Macdonald, T. L.; Ocain, T. D.; Longhi, D. *J. Org. Chem.* **1993**, *58*, 6609−6613.

(442) Babine, R. E.; Bleckman, T. M.; Kissinger, C. R.; Showalter, R.; Pelletier, L. A.; Lewis, C.; Tucker, K.; Moomaw, E.; Parge, H. E.; Villafranca, J. E. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1719−1724.

(443) Bohm, H. J. *J. Comput. Aided Mol. Des.* **1992**, *6*, 69.

(444) Goulet, M. T.; Rupprecht, K. M.; Sinclair, P. J.; Wyvratt, M. J.; Parsons, W. H. *Perspect. Drug Discovery Des.* **1994**, *2*, 145−162.

(445) Dumont, F. J.; Staruch, M. J.; Koprak, S. L.; Siekierka, J. J.; Lin, C. S.; Harrison, R.; Sewell, T.; Kindt, V. M.; Beattie, T. R.; et al. *J. Exp. Med.* **1992**, *176*, 751−760.

(446) Becker, J. W.; Rotonda, J.; McKeever, B. M.; Chan, H. K.; Marcy, A. I.; Wiederrecht, G.; Hermes, J. D.; Springer, J. P. *J. Biol. Chem.* **1993**, *268*, 11335−11339.

(447) Andrus, M. B.; Schreiber, S. L. *J. Am. Chem. Soc.* **1993**, *115*, 10420−10421.

(448) Furber, M. *J. Am. Chem. Soc.* **1995**, *117*, 7267−7268.

(449) Teague, S. J.; Stocks, M. J. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1947−1950.

(450) Dragovich, P.; Parge, H.; Kissinger, C.; Knighton, D.; Lewis, C. Unpublished results.

(451) Babine, R.; Parge, H.; Dragovich, P.; Kissinger, C.; Tatlock, J.; Lewis, C.; Kalish, V.; Villafranca, J. Unpublished results.

(452) Blokzijl, W.; Enberts, J. B. F. N. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1545−1579.

(453) Ikeda, Y.; Schultz, L. W.; Clardy, J.; Schreiber, S. L. *J. Am. Chem. Soc.* **1994**, *116*, 4143−4144.

(454) Babine, R. E.; Bleckman, T. M.; Lewis, C. T. Unpublished results.

(455) Babine, R. E.; Bleckman, T. M.; Littlefield, E. S.; Parge, H. E.; Pelletier, L. A. K.; Lewis, C. T.; French, J. V.; Imbacuan, M.; Katoh, S.; Tatlock, J. H.; Showalter, R. E.; Villafranca, J. E. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 385−390.

(456) Babine, R.; Littlefield, E.; Parge, H.; Kissinger, C.; Lewis, C. Unpublished results.

(457) Griffith, J. P.; Kim, J. L.; Kim, E. E.; Sintchak, M. D.; Thomson, J. A.; Fitzgibbon, M. J.; Fleming, M. A.; Caron, P. R.; Hsiao, K.; Navia, M. A. *Cell* **1995**, *82*, 507−522.

(458) Kissinger, C. R.; Parge, H. E.; Knighton, D. R.; Lewis, C. T.; Pelletier, L. A.; Tempczyk, A.; Kalish, V. J.; Tucker, K. D.; Showalter, R. E.; Moomaw, E. W.; Gastinel, L. N.; Habuka, N.; Chen, X.; Maldonado, F.; Barker, J. E.; Bacquet, R.; Villafranca, E. *Nature* **1995**, *378*, 641−644.

(459) Mierke, D. F.; Schmeider, P.; Karuso, P.; Kessler, H. *Helv. Chim. Acta* **1991**, *74*, 1027−1047.

(460) Petros, A. M.; Luly, J. R.; Liang, H.; Fesik, S. W. *J. Am. Chem. Soc.* **1993**, *115*, 9920−9924.

(461) Smith, P. W.; Still, W. C. *J. Am. Chem. Soc.* **1988**, *110*, 7917−7919.

(462) Itoh, S.; DeCenzo, M. T.; Livingston, D. J.; Pearlman, D. A.; Navia, M. A. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1983−1988.

(463) Tang, S.; Still, W. C. *Tetrahedron Lett.* **1993**, *34*, 6701−6704.

(464) Lepre, C. A.; Cheng, J. W.; Moore, J. M. *J. Am. Chem. Soc.* **1993**, *115*, 4929−30.

(465) Cheng, J.-W.; Lepre, C. A.; Moore, J. M. *Biochemistry* **1994**, *33*, 4093−4100.

(466) Itoh, S.; Navia, M. A. *Protein Sci.* **1995**, *4*, 2261−2268.
(467) Choi, J.; Chen, J.; Schreiber, S. L.; Clardy, J. *Science* **1996**, *273*, 239−242.
(468) Goldberg, J.; Huang, H.-b.; Kwon, Y.-g.; Greengard, P.; Nairn, A. C.; Kuriyan, J. *Nature* **1995**, *376*, 745−753.
(469) Hashimoto, Y.; Perrino, B. A.; Soderling, S. R. *J. Biol. Chem.* **1990**, *265*, 1924−1927.
(470) Rivetna, M. N.; Salowe, S. P.; Tolman, R. L.; Jones, A. B. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1147−1150.
(471) Spencer, D. M.; Wandless, T. J.; Schreiber, S. L.; Crabtree, G. R. *Science* **1993**, *262*, 1019−1024.
(472) Appling, D. R. *FASEB J.* **1991**, *5*, 2645.
(473) Berman, E. M.; Werbel, L. M. *J. Med. Chem.* **1991**, *34*, 479.
(474) Varney, M. D.; Romines, W. H. *Curr. Opin. Ther. Pat.* **1992**, 1979−1991.
(475) Seeger, D. R.; Cosulich, D. B.; Smith, J. M.; Hulquist, M. E. *J. Am. Chem. Soc.* **1949**, *71*, 1753.
(476) Matthews, D. A.; Bolin, J. T.; Burridge, J. M.; Filman, D. J.; Volz, K. W.; Kraut, J. *J. Biol. Chem.* **1985**, *260*, 392.
(477) Kuyper, L. F.; Roth, B.; Baccanari, D. P.; Ferone, R.; Beddell, C. R.; Champness, J. N.; Stammers, D. K.; Dann, J. G.; Norington, F. E.; Baker, D. J.; Goodford, P. J. *J. Med. Chem.* **1985**, *28*, 303.
(478) Freisheim, J. H.; Matthews, D. A. In *Folate Antagonists as Therapeutic Agents*; Sirotnak, F. M., Burchall, J. J., Ensminger, W. D., Montgomery, J. A., Eds.; Academic Press, Inc.: Orlando, 1984, pp 69−131.
(479) Blakely, R. L. In *Folates and Pterins*; Blakley, R. L., Benkovic, S. J., Eds.; John Wiley & Sons: New York, 1984; Vol. 1, pp 191−253.
(480) Kraut, J.; Matthews, D. A. In *Biological Macromolecules and Assemblies*; Jurnak, F. A., McPherson, A., Eds.; John Wiley & Sons: New York, 1987; Vol. 3, pp 1−72.
(481) Kuyper, L. F. In *Computer-Aided Drug Design*; Perun, T. J., Propst, C. L., Eds.; Marcel Decker, Inc.: New York, 1989; pp 327−369.
(482) Bolin, J. T.; Filman, D. J.; Matthews, D. A.; Hamlin, R. C.; Kraut, J. *J. Biol. Chem.* **1982**, *257*, 13650.
(483) Bystroff, C.; Oatley, S. J.; Kraut, J. *Biochemistry* **1990**, *29*, 3263.
(484) Matthews, D. A.; Villafranca, J. E.; Janson, C. A.; Smith, W. W.; Welsh, K.; Freer, S. *J. Mol. Biol.* **1990**, *214*, 937−948.
(485) Reich, S. H.; Webber, S. E. *Perspect. Drug Discovery Des.* **1993**, *1*, 371−390.
(486) Jones, T. R.; Calvert, A. H.; Jackman, A. L.; Brown, S. J.; Jones, M.; Harrap, K. R. *Eur. J. Cancer* **1981**, *17*, 11.
(487) Montfort, W. R.; Perry, K. M.; Fauman, E. B.; Finer-Moore, J. S.; Maley, G. F.; Hardy, L.; Stroud, R. M. *Biochemistry* **1990**, *29*, 6964−6977.
(488) Webber, S. E.; Bleckman, T. M.; Attard, J.; Deal, J. G.; Kathardeker, V.; Welsh, K. M.; Webber, S.; Janson, C. A.; Matthews, D. A.; Smith, W. W.; Freer, S. T.; Jordan, S. R.; Bacquet, R. J.; Howland, E. F.; Booth, C. L. J.; Ward, R. W.; Hermann, S. M.; White, J.; Morse, C. A.; Hilliard, J. A.; Bartlett, C. A. *J. Med. Chem.* **1993**, *36*, 733−746.
(489) Varney, M. D.; Marzonni, 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.; Herrmann, S. M.; Howland, E. F.; Ward, R. W.; White, J. *J. Med. Chem.* **1992**, *35*, 663−676.
(490) Reich, S. H.; Fuhry, M. A. M.; Nguyen, D.; Pino, M. J.; Welsh, K. M.; Webber, S.; Janson, C. A.; Jordan, S. R.; Matthews, D. A.; Smith, W. W.; Bartlett, C. A.; Booth, C. L. J.; Herrmann, S. M.; Howland, E. F.; Morse, C. A.; Ward, R. W.; White, J. *J. Med. Chem.* **1992**, *35*, 847−858.
(491) Baldwin, S. W.; Tse, A.; Gossett, L. S.; Taylor, E. C.; Rosowsky, A.; Shih, C.; Moran, R. G. *Biochemistry* **1991**, *30*, 1997−2006.
(492) Taylor, E. C.; Harrington, R. J.; Fletcher, S. R.; Beardsley, G. P.; Moran, R. G. *J. Med. Chem.* **1985**, *28*, 914−921.
(493) Beardsley, G. P.; Moroson, B. A.; Taylor, E. C.; Moran, R. G. *J. Biol. Chem.* **1989**, *264*, 328−333.
(494) Almassy, R. J.; Janson, C. A.; Kan, C.-C.; Hostomska, Z. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 6114−6118.
(495) Maren, T. H. *Mol. Pharm.* **1992**, *41*, 419−426.
(496) Baldwin, J. J.; Smith, G.; Springer, J. P.; Murcko, M. A. *Chem. Des. Automation News* **1992**, *7*, 1−34.
(497) 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. *J. Med. Chem.* **1989**, *32*, 2510−2513.
(498) Smith, G. M.; Alexander, R. S.; Christianson, D. W.; McKeever, B. M.; Ponticello, G. S.; Springer, J. P.; Randall, W. C.; Baldwin, J. J.; Habecker, C. N. *Protein Sci.* **1994**, *3*, 118−125.
(499) Eriksson, A. E.; Jones, T. A.; Liljas, A. *Proteins: Struct. Funct. Genet.* **1988**, *4*, 274−282.
(500) Nevalainen, T. *Clin. Chem.* **1993**, *39*, 2453−2459.
(501) Scott, D. L.; White, S. P.; Otwinowski, Z.; Yuan, W.; Gelb, M. H.; Siegler, P. B. *Science* **1990**, *250*, *1541*.
(502) Scott, D. L.; White, B., J. L.; Rosa, J. J.; Gelb, M. H.; Sigler, P. B. *Science* **1991**, *254*, 1007.
(503) Schevitz, R. W.; Bach, N. J.; Carlson, D. G.; Chirgadze, D. K.; Clawson, R. D.; Dillard, S. E.; Draheim, S. E.; Hartley, L. W.; Jones, N. D.; Milhich, E. E.; Olkowski, J. L.; Snyder, D. W.; Sommers, C.; Wery, J.-P. *Nature Struct. Biol.* **1995**, *2*, 458−465.
(504) Ripka, W. C.; Sipio, W. J.; Blaney, J. M. *J. Heterocycl. Chem.* **1987**, *24*, S95−S104.
(505) Wilson, I. A.; Skehel, J. J.; Wiley, D. C. *Nature* **1981**, *289*, 366−373.
(506) Wiley, D. C.; Wilson, I. A.; Skehel, J. J. *Nature* **1981**, *289*, 373−378.
(507) Varughese, J. N.; Laver, W. G.; Colman, P. M. *Nature* **1983**, *303*, 35−40.
(508) Colman, P. M.; Varghese, J. N.; Laver, W. G. *Nature* **1983**, *303*, 41−44.
(509) Palese, P.; Jobita, K.; Ueda, M.; Compans, R. W. *Virology* **1974**, *61*, 397−410.
(510) Burmeister, W. P.; Ruigrok, R. W. H.; Cusack, S. *EMBO J.* **1992**, *11*, 49−56.
(511) Itzstein, M. V.; Wu, W.-Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Phan, T. V.; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, P. M.; Varghese, J. N.; Ryan, D. M.; Woods, J. M.; Bethel, R. C.; Hotham, V. J.; Cameron, J. M.; Penn, C. R. *Nature* **1993**, *363*, 418−423.
(512) Goodford, P. J. *J. Med. Chem.* **1985**, *28*, 849−857.
(513) Montgomery, J. A. *Med. Res. Rev.* **1993**, *13*, 209.
(514) Kuriyan, J.; Cowburn, D. *Curr. Opin. Struct. Biol.* **1993**, *3*, 828−837.
(515) Choong, I. C.; Ellman, J. A. *Annu. Rep. Med. Chem.* **1996**, *31*, 309−318.
(516) Chen, J. K.; Schreiber, S. L. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 953−969.
(517) Chen, J. K.; Lane, W. S.; Brauer, A. W.; Tanaka, A.; Schreiber, S. L. *J. Am. Chem. Soc.* **1993**, *115*, 12591−12592.
(518) Feng, S.; Chen, J. K.; Yu, H.; Simon, J. A.; Schreiber, S. L. *Science* **1994**, *266*, 1241−1247.
(519) Feng, S.; Kasahara, C.; Rickles, R. J.; Schreiber, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 12408−12415.
(520) Combs, A. P.; Kapoor, T. M.; Feng, S.; Chen, J. K.; Daude-Snow, L. F.; Schreiber, S. L. *J. Am. Chem. Soc.* **1996**, *118*, 287−288.
(521) Ohlmeyer, M. H. J.; Swanson, R. N.; Dillard, L. W.; Reader, J. C.; Asouline, G.; Kobayashi, R.; Wigler, M.; Still, W. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10922−10926.
(522) Feng, S.; Kapoor, T. M.; Shirai, F.; Combs, A. P.; Schreiber, S. L. *Chem. Biol.* **1996**, *3*, 661−670.
(523) Weber, P. C.; Ohlendorf, D. H.; Wendoloski, J. J.; Salemme, F. R. *Science* **1989**, *243*, 85−88.
(524) Weber, P. C.; Wendoloski, J. J.; Pantoliano, M. W.; Salemme, F. R. *J. Am. Chem. Soc.* **1992**, *114*, 3197−3200.
(525) Williams, D. H.; Westwell, M. S. *Chem. Biol.* **1996**, *3*, 695−701.
(526) Taylor, R.; Kennard, O. *Acc. Chem. Res.* **1984**, *17*, 320.
(527) Pimental, G. C.; McClellan, A. L. *The Hydrogen Bond*; Freeman & Co.: London, 1960.
(528) Mitchel, J. B. O.; Price, S. L. *J. Comput. Chem.* **1990**, *11*, 1217−1233.
(529) Reed, A. E.; Weinhold, F.; Curtiss, L. A.; Pochatko, D. J. *J. Chem. Phys.* **1986**, *84*, 5687−5703.
(530) Cybulski, S. M.; Scheiner, S. *J. Phys. Chem.* **1989**, *93*, 6565−6574.
(531) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97.
(532) Morgan, B. P.; Scholtz, J. M.; Ballinger, M. D.; Zipkin, I. D.; Bartlett, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 297−307.
(533) Tronrud, D. E.; Holden, H. M.; Matthews, B. W. *Science* **1987**, *235*, 571−574.
(534) Kollman, P. A.; Mertz, K. M., Jr. *Acc. Chem. Res.* **1990**, *23*, 246−252.
(535) Gelb, M. H.; Svaren, J. P.; Abeles, R. H. *Biochemistry* **1985**, *24*, 1813−1816.
(536) Burkey, T. J.; Fahey, R. C. *J. Am. Chem. Soc.* **1983**, *105*, 868−871.
(537) Morgan, B. P.; Holland, D. R.; Matthews, B. W.; Bartlett, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 3251−3260.
(538) Kofron, J. L.; Kuzmic, P.; Kishore, V.; Gemmecker, G.; Fesik, S. W.; Rich, D. H. *J. Am. Chem. Soc.* **1992**, *114*, 2670−2675.

CR960370Z