



# Molecular Modeling Studies of Novel Retro-binding Tripeptide Active-site Inhibitors of Thrombin

Wan F. Lau,<sup>a\*</sup> Lydia Taberero,<sup>b</sup> John S. Sack<sup>b</sup> and Edwin J. Iwanowicz<sup>c</sup>

<sup>a</sup>Computer-Assisted Drug Design, <sup>b</sup>Macromolecular Crystallography and <sup>c</sup>Cardiovascular Chemistry, Bristol-Myers Squibb Pharmaceutical Research Institute, P.O. Box 4000, Princeton, NJ 08543-4000, U.S.A.

**Abstract**—A novel series of retro-binding tripeptide thrombin active-site inhibitors was recently developed (Iwanowicz, E. I. *et al. J. Med. Chem.* **1994**, *37*, 2111<sup>1</sup>). It was hypothesized that the binding mode for these inhibitors is similar to that of the first three N-terminal residues of hirudin. This binding hypothesis was subsequently verified when the crystal structure of a member of this series, BMS-183,507 (*N*[*N*[*N*-[4-(Aminoiminomethyl)amino]-1-oxobutyl]-*L*-phenylalanyl]-*L*-allo-threonyl]-*L*-phenylalanine, methyl ester), was determined (Taberero, L. *J. Mol. Biol.* **1995**, *246*, 14<sup>2</sup>). The methodology for developing the binding models of these inhibitors, the structure–activity relationships (SAR) and modeling studies that led to the elucidation of the proposed binding mode is described. The crystal structure of BMS-183,507/human  $\alpha$ -thrombin is compared with the crystal structure of hirudin/human  $\alpha$ -thrombin (Rydell, T. J. *et al. Science* **1990**, *249*, 227;<sup>3</sup> Rydel, T. J. *et al. J. Mol. Biol.* **1991**, *221*, 583;<sup>4</sup> Grutter, M. G. *et al. EMBO J.* **1990**, *9*, 2361<sup>5</sup>) and with the computational binding model of BMS-183,507.

## Introduction

Thrombin is a multifunctional trypsin-like serine protease that plays a central role in hemostasis and thrombosis.<sup>6–9</sup> Thrombin is the final key enzyme in the coagulation cascade, cleaving fibrinogen to generate fibrin, which then polymerizes to form a hemostatic plug. It also activates factor XIII, which cross-links and stabilizes the fibrin polymer. Thrombin is an important pharmaceutical target for the treatment and prevention of various disease processes including arterial and venous thrombosis and arteriosclerosis. Direct-acting thrombin inhibitors that inhibit the proteolytic activity of thrombin offer the potential for a simplified usage as well as broader efficacy over existing agents like heparins and coumarin (warfarin).

The crystal structures of the inhibitor–thrombin complexes for PPACK,<sup>10–12</sup> NAPAP,<sup>12</sup> MD-805,<sup>12</sup> BMS-183,507,<sup>12</sup> cyclotheonamide A,<sup>13</sup> and hirudin<sup>3,4,5</sup> have been determined. BMS-183,507 (Fig. 1) is a member of a series of tripeptide inhibitors of thrombin that was recently developed from SAR and modeling studies of compound 1 (Table 1).<sup>1</sup> These inhibitors, as exemplified by BMS-183,507 ( $K_i = 17.2$  nM),<sup>2</sup> are highly potent and selective inhibitors of thrombin. The common binding features at the active site region for PPACK, NAPAP, MD-805, BMS-183,507 and the first three N-terminus residues of hirudin are summarized schematically in Figure 1.

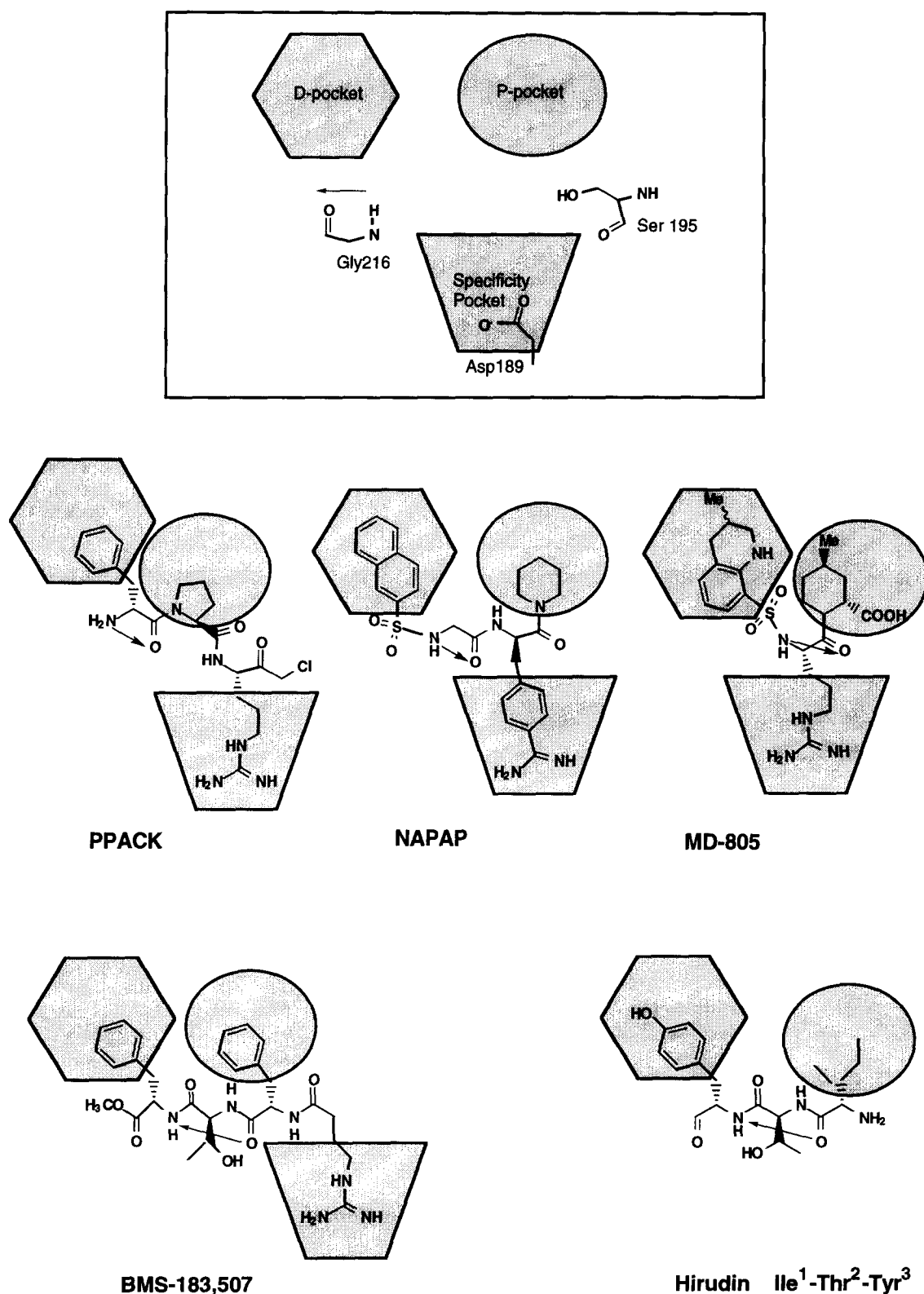
PPACK, NAPAP and MD-805 have hydrophobic moieties that bind in the hydrophobic 'P-pocket' and the 'D-

pocket',<sup>14</sup> and a guanidinium moiety that binds in the specificity pocket and forms a salt bridge with the carboxyl of Asp189. The backbone of each inhibitor juxtaposes the extended thrombin segment Ser214–Glu217 in an anti-parallel manner which allows for the formation of a pair of hydrogen-bonds to Gly216.<sup>14</sup> Anti-parallel  $\beta$ -sheet hydrogen bonds to Gly216 (or the equivalent residue) are often involved in the interaction of serine proteases with peptide substrates or substrate-derived inhibitors.<sup>15,16</sup> The crystal structure of residues 7–16 of the A $\alpha$ -chain of human fibrinogen bound to bovine thrombin shows a similar interaction.<sup>17</sup>

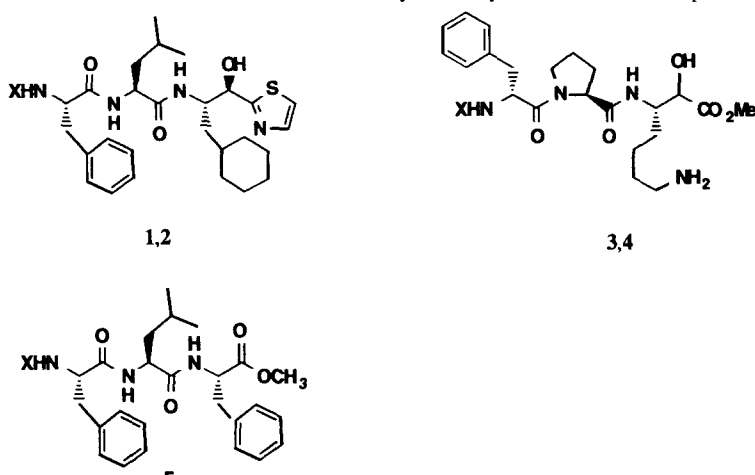
The first three N-terminal residues of hirudin also possess two hydrophobic moieties that bind in the P- and D-pockets, and the main chain forms a pair of hydrogen bonds to Gly216 (Fig. 1). However, hirudin aligns its main chain parallel to the Ser214–Glu217 segment. Hirudin also lacks the basic functionality that binds in the specificity pocket.

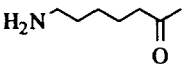
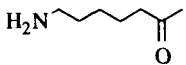
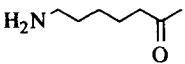
BMS-183,507 binds with its alkyl-guanidino or amino moiety in the specificity pocket of thrombin. Like the N-terminal of hirudin, the main chain of the inhibitor is aligned parallel to Ser214–Gly216 of thrombin. Accordingly, the direction of the main chain is reversed relative to other known substrates and small molecule inhibitors of thrombin, e.g., PPACK, and also to other substrates of thrombin (Fig. 1). Presumably, the other analogs of the series also bind like BMS-183,507. Since this reversal of the backbone peptide bonds is not coupled to inversion of each chiral center, BMS-183,507 and its congeners are not true retro-inverso peptides<sup>18</sup> (Fig. 2). We describe these peptides that bind in this reversed inhibitory mode as retro-binding peptides.<sup>1,19</sup>

<sup>1</sup>BMS-183,507 is: *N*[*N*[*N*-[4-(Aminoiminomethyl)amino]-1-oxobutyl]-*L*-phenylalanyl]-*L*-allo-threonyl]-*L*-phenylalanine, methyl ester.

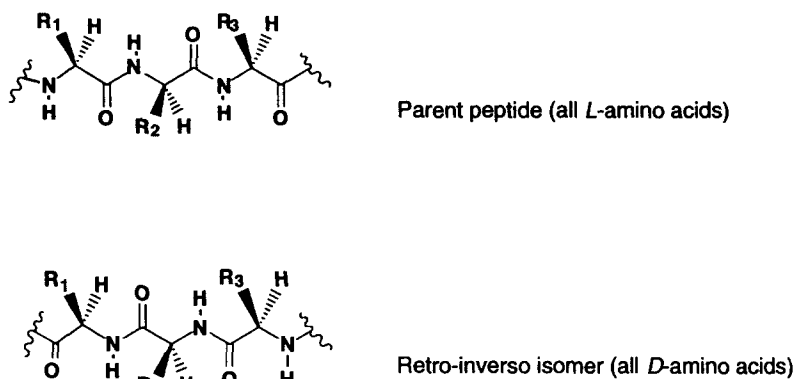


**Figure 1.** Key binding functionalities of thrombin active site. Top: schematic diagram of thrombin active site showing key binding features; hydrophobic P-pocket and D-pocket, the specificity pocket with Asp189 that forms salt-bridges or strong hydrogen bonds to basic and/or hydrogen bonding moieties, and Gly216 that forms a pair of hydrogen bonds to the backbone of the inhibitors. Arrow indicates the direction of the chain segment of thrombin. Below: inhibitors that bind to the active site, shown here in the orientation relative to the diagram. The moieties that bind in the P-pocket, the D-pocket, and the specificity pocket are enclosed in hexagonal, circular and trapezoid boxes, respectively. Arrows indicate the atoms on the inhibitor that form the pair of hydrogen bonds to Gly216 and the direction of the backbone of the inhibitors.

**Table 1.** *In vitro* inhibition of thrombin catalytic activity for **1** and related compounds


compd	X	IC <sub>50</sub> (μM)*
1		8
2	H	>530
3		18
4	H	20-30
5		80

\**In vitro* inhibition of thrombin catalytic activity on 10 mM substrate S-2238 (D-Phe-Pip-Arg-pNA).<sup>1</sup>

**Figure 2.** Parent all *L*-peptide and its retro-inverso isomer

Prior to the solution of the crystal structure of BMS-183,507–human  $\alpha$ -thrombin, this retro-binding mode had been proposed based on computational binding models for these inhibitors generated by docking studies. It was also hypothesized from the computational binding models of BMS-183,507 and other inhibitors in the series, that these inhibitors bind like the first three residues of the N-terminus of hirudin.<sup>1</sup> The inhibitors form a  $\beta$ -parallel hydrogen bond to Gly216 with Phe1 O and Phe3 N, and the amino acid side chains are oriented in regions similar to the corresponding three residues in hirudin. A major

structural difference between these inhibitors and hirudin is the presence of the alkyl-guanidino moiety in these inhibitors. This moiety binds in the specificity pocket. The binding hypothesis was subsequently verified with the crystal structure of BMS-183,507 bound to human  $\alpha$ -thrombin. Although the inhibitors bind like the N-terminus of hirudin, acylation of the N-terminus of hirudin with a glycyl residue, an acetyl or an acetimidyl moiety reduced binding in hirudin.<sup>20,21</sup> In contrast, acylation of the retro-peptides with an alkyl-guanidino or alkyl-amino moiety is required for thrombin active-site inhibition.<sup>1</sup>

This paper describes in further details the methodology for developing the binding models, the SAR and the modeling studies that led to the elucidation of the proposed binding mode, and the crystal structure of BMS-183,507- $\alpha$ -thrombin. The crystal structure of BMS-183,507-human  $\alpha$ -thrombin is compared with the crystal structure of hirudin (rHV2)-human  $\alpha$ -thrombin (4HTC),<sup>4</sup> and with the computational binding model of BMS-183,507. These comparisons offer new insights into the structural requirements for binding in the specificity pocket and in the vicinity of Ser195, and into the variations between the experimental and the computational models.

### Development of Computational Binding Models

#### Methods

The MacroModel/BATCHMIN<sup>22</sup> software package employing the AMBER united atom force field<sup>23</sup> was used for all minimizations and for Monte Carlo/Energy Minimization (MC/EM).<sup>24</sup> These minimizations employed conjugate gradient minimization which terminated after 1000 iterations or when the energy gradient rms fell below  $0.1 \text{ kJ } \text{\AA}^{-1}$ . A dielectric function of  $D = 2r$ ,<sup>25</sup> an electrostatic cutoff of  $12 \text{ \AA}$ , a van der Waals cutoff of  $7 \text{ \AA}$  and a hydrogen bonding cutoff of  $4 \text{ \AA}$  were utilized.

Coordinates for thrombin were taken from the crystal structure of PPACK-thrombin.<sup>10</sup> The complex structure was subjected to 1000 energy minimization iterations with the backbone atoms constrained by a harmonic force constant of  $100 \text{ kJ } \text{\AA}^{-1}$ , followed by another 500 energy minimization iterations without constraint. Structures were visualized using SYBYL<sup>26</sup> running on a Silicon Graphics 420VGX.

#### Docking with DGEOM/energy minimization (DG/EM)

Models of the inhibitors 1, 6 and 7 were first constructed

using SYBYL. Each starting structure was then geometry optimized using BATCHMIN. Following this, they were docked with DGEOM<sup>27,28,29</sup> using two sets of hydrogen bond distance constraints (Fig. 3). For example, 1 was docked using the set of hydrogen-bonds distance constraints between the basic nitrogens and the carboxyl oxygens of Asp189 and the set of hydrogen bonds distance constraints between the cyclohexyl-Ala3 N and Leu2 O. Fifty docked conformers were randomly generated for each inhibitor. Each of the 50 conformers were then subjected to energy minimization where the protein residues were constrained by a harmonic constraint of  $100 \text{ kJ } \text{\AA}^{-2}$  but the ligand was allowed to minimize without any constraints. The model for the bound conformer was chosen from the lowest energy structure. The lowest energy structure was found to be qualitatively similar for all three inhibitors.

This DG/EM protocol was tested earlier by docking D-Phe-Pro-Arg-aldehyde. Using two similar sets of hydrogen-bond constraints for docking and selecting the lowest energy structure, a reasonable binding model for D-Phe-Pro-Arg-aldehyde (RMSD of  $0.798 \text{ \AA}$  with PPACK for all equivalent heavy atoms, excluding the carbonyl of the aldehyde) was obtained.

#### Monte Carlo/energy minimization (MC/EM) studies of 7

To determine whether significant structural similarity exists between the lowest energy conformer of 7 and the remaining 49 conformers, cluster analysis using the complete linkage hierarchical cluster procedure<sup>28,30</sup> was used to group the 50 ligand conformers into different conformational families based on the RMS difference of their coordinates. The three low energy conformers clustered into three different clusters, while the remaining 47 conformers which were more than  $50 \text{ kJ mol}^{-1}$

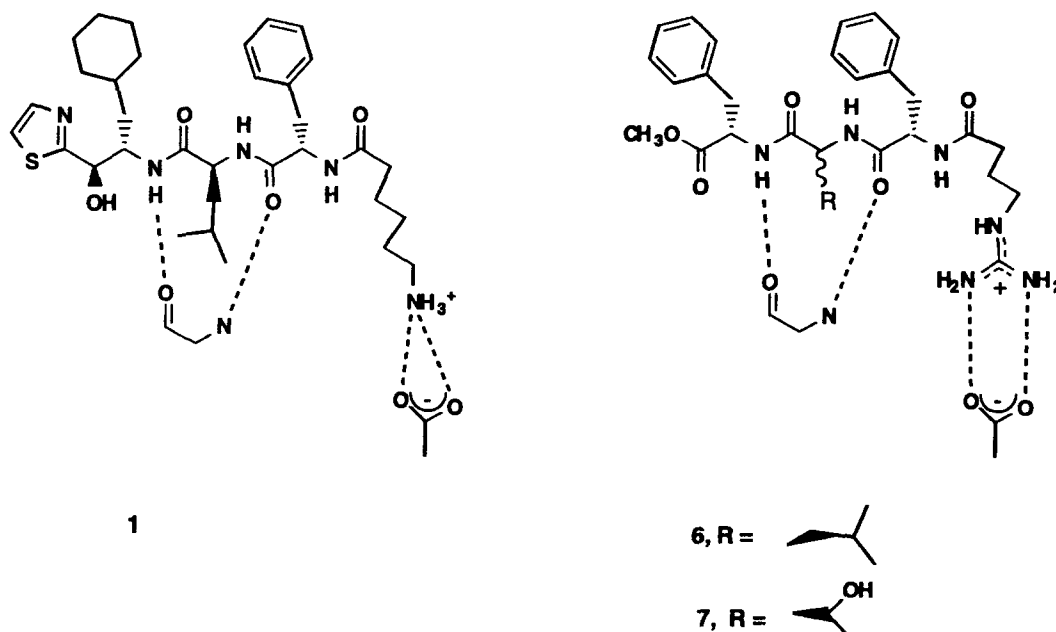


Figure 3. DGEOM hydrogen-bond distance constraints. A lower bound of  $2.7 \text{ \AA}$  and an upper bound of  $3.2 \text{ \AA}$  was employed.

higher in energy than the lowest energy conformer clustered into 11 clusters. The backbones of the conformers of the three lowest energy structures were similar, they differed only in the placement of the side chains of the phenyl-alanine residues. However, the backbones of the conformers from the 11 clusters all differed significantly from the backbones of the three conformers.

To search for other low energy backbone binding conformations, MC/EM employing a usage directed conformational search<sup>31</sup> was carried out in two stages. The lowest energy structure (DG/EM structure, Table 2) was used as the starting structure for the first stage. The inhibitor was unconstrained during minimization, but the protein atoms were constrained by a harmonic force constant of 100 kJ Å<sup>-1</sup>. A total of 2000 MC steps were used. All rotatable bonds between heavy atoms were allowed to vary during each MC step. Each MC step was followed by energy minimization, and conformers that differed from the global minimum by no more than 20 kJ mol<sup>-1</sup> were saved. This stage generated 66 structures. Cluster analysis of these 66 structures showed that they belong to two clusters which are essentially similar to the two lowest energy structures from the DG/EM structures. Hence, this study did not find any other low energy backbone conformation for the inhibitor, although it cannot be claimed that an exhaustive search was performed.

The lowest energy structure (MC/EMI, Fig. 4) was then subjected to the next MC/EM stage. The inhibitor and all residues within 4 Å of the inhibitor were unconstrained during minimization, but the remaining residues of the protein were all constrained by a harmonic force constant of 100 kJ Å<sup>-1</sup>. Only the  $\chi_1$  angle of the *allo*-threonine

**Table 2.** RMSD<sup>a</sup> of atoms of BMS-183,507 from the computational binding models relative to atoms of BMS-183,507 from the crystal structure of BMS-183,507-human  $\alpha$ -thrombin<sup>b</sup>

Conformer	Main Chain <sup>c</sup>	All atoms
DG/EM	0.662	1.415
MC/EM I	0.741	0.964
MC/EM II	0.803	1.148

<sup>a</sup>Angstroms.

<sup>b</sup>Lowest energy conformer of BMS-183,507 from DG/EM, first stage MC/EM I, second stage MC/EMI.

<sup>c</sup>Main Chain atoms being C $\alpha$ , C $\beta$ , C, N and carbonyl moieties of ester and acyl-alkyl chain.

side-chain was varied, and 200 MC steps were taken. Each MC step was followed by energy minimization, and conformers within 50 kJ of the global minimum were kept. This stage generated only three different conformers, with  $\chi_1$  of 62° for the *allo*-threonine side-chain in the lowest energy structure (MC/EM II structure), and  $\chi_1$  of -52° and 167°, respectively for the other two rotamers which were 11 kJ mol<sup>-1</sup> and 31 kJ mol<sup>-1</sup> higher in energy relative to the MC/EM II structure. The hydroxyl of *allo*-threonine forms a hydrogen bond to the carboxyl of Glu192 in the two *gauche*-rotamers, but forms a hydrogen bond to the Gly219 N in the case of the *trans*-rotamer.

#### Hydrogen-bond distance constraints for docking

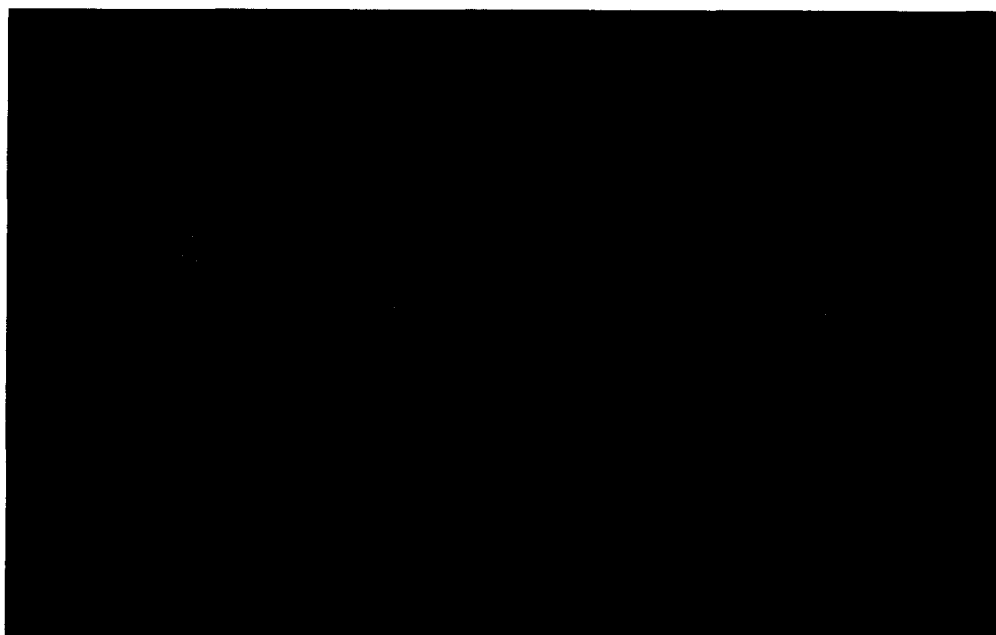
The set of hydrogen-bonds distance constraints between the basic nitrogens and the carboxyl oxygens of Asp189 (Fig. 3) is based on the assumption that the basic terminus binds in the specificity pocket. This interaction with the Asp189 carboxyl is often seen in inhibitors that do not form covalent intermediates with the enzyme.<sup>6</sup> The other set of distance constraints with Gly216 is based on the observation that the known thrombin inhibitors maintain



**Figure 4.** Comparison of structure of BMS-183,507 in crystal structure of BMS-183,507- $\alpha$ -thrombin complex with computational models: lowest energy structures at various stages, DG/EM (cyan), MC/EM I (magenta) and MC/EM II (green) are super-positioned onto crystal structure (white).



**Figure 5.** Comparison of BMS-183,507 with hirudin (rHV2) Ile1-Thr2-Tyr3. Crystal structure of hirudin (rHV2)-human  $\alpha$ -thrombin (4HTC) superimposed onto crystal structure of BMS-183,507-human  $\alpha$ -thrombin. Hirudin residues Ile1-Thr2-Tyr3 carbon atoms are colored cyan, BMS-183,507 carbon atoms are colored white. Thrombin of BMS-183,507- $\alpha$ -thrombin complex shown in yellow, thrombin of hirudin- $\alpha$ -thrombin complex shown in green.



**Figure 6.** Comparison of BMS-183,507 with hirudin (rHV2) Ile1-Thr2-Tyr3: top view. Crystal structure of hirudin (rHV2)-human  $\alpha$ -thrombin (4HTC) superimposed onto crystal structure of BMS-183,507-human  $\alpha$ -thrombin. Hirudin residues Ile1-Thr2-Tyr3 carbon atoms are colored cyan, BMS-183,507 carbon atoms are colored white. Thrombin of BMS-183,507- $\alpha$ -thrombin complex shown in yellow, thrombin of hirudin- $\alpha$ -thrombin complex shown in green.

a pair of hydrogen bonds to Gly216.<sup>3,15,16</sup> Taking **1** as an example, the choice of the cyclohexyl-Ala 3 (cHex-Ala3) N and Leu2 O as the hydrogen-bonding pair was based on the following geometrical considerations.

Given the orientation of Gly216 relative to Asp189, the

hydrogen-bond acceptor must precede a hydrogen-bond donor on the ligand. Assuming that no *cis*-amide would act as the hydrogen-bonding pair to Gly216, this gives only three pairs of atoms on **1** that might possibly form hydrogen bonds to Gly216: 1) Leu2 N and 6-amino-hexanoyl O; 2) cHex-Ala3 N and Leu2 O; 3) cHex-Ala3 O

and Leu2 O. Of these, the second atom pair, cHex-Ala3 N and Leu2 O, was the best choice for forming the hydrogen bonds to Gly216. Also, this hydrogen bonding scheme where the hydrogen bonding donors and acceptors come from the amides of adjacent residues is seen in parallel  $\beta$ -sheets. The other two pairs are unlikely based on the arguments below.

The first set is unlikely because the carbonyl group in the 6-aminohexanoyl moiety is separated from the terminal amine by only six bonds. If the Leu2 N and 6-aminohexanoyl O were to form hydrogen bonds to Gly216, it would be difficult to dock the amine sufficiently close to Asp189 to form a strong hydrogen bond. It would also be difficult to dock the Phe side chain into the P-pocket. Conversely, the third set of atoms are unlikely because the carbonyl group of Leu2 is separated from the terminal amine by twelve bonds. This makes it more difficult to dock the Leu and Phe residues into the limited space between Gly216 and the active site. As a test, **1** was also docked using these two sets of distance constraints: a) 6-aminohexanoyl amine N to Asp189 carboxyl oxygens, Leu2 N to Gly216 O and 6-amino-hexanoyl O to Gly216 N; and b) 6-aminohexanoyl amine N to Asp189 carboxyl oxygens, cHex-Ala3 O to Gly216 O and Leu2 O to Gly216 N. A similar DG/EM protocol was used; 50 random conformations followed by energy minimization. The best DG/EM structure was obtained using cHex-Ala3 N and Leu2 O to form the pair of hydrogen bonds to Gly216. Subsequently, the tolerated replacement of the hydroxyl by an ester carbonyl O (Table 1) argued against the hydroxyl being the hydrogen bond donor to Gly216 O.

### Experimental Methods

The synthesis and biological studies of the tripeptides **5**, **6** and **7** is given in Ref. 1. The experimental details for crystallization and data collection are given in Ref. 2. The structure was determined by the method of molecular replacement using the program MERLOT<sup>32</sup> where the coordinates for the refined human thrombin /PPACK structure<sup>11</sup> were used for the initial model. The coordinates for hirugen were extracted from the C-terminal region of the hirudin structure<sup>4</sup> and imposed onto the thrombin model using CHAIN.<sup>33</sup> The BMS-183,507 computational model described above was used for the inhibitor. The structure was refined by the method of simulated annealing using XPLOR.<sup>34</sup> The final R-value was 20.4% for all the data between 6 Å and 2.6 Å resolution, with an rms deviation of 0.014 Å for bond distances and 2.0° for bond angles.

### Lead Identification, SAR Studies and Test of Hypothesis

Compounds **1** and **2** (Table 1) were discovered from testing of compounds identified from substructure searching of in-house compounds.<sup>1</sup> The 6-aminohexanoyl moiety appeared to be essential for the potency of com-

pound **1**. In contrast, the 6-aminohexanoyl moiety did not enhance the potency of **3** relative to **4**, a different structural class of compounds which we assume bind in a fashion similar to PPACK.<sup>35</sup>

This observation suggests that **1** might bind differently from **3** and **4**. Instead of binding with its peptide backbone oriented in the same direction as **4** and the hydroxy-thioimidazole functioning as an isostere of the hydroxy-ester of **4**,<sup>36</sup> we hypothesized that **1** binds in another mode with the 6-aminohexanoyl moiety binding in the specificity pocket. Docking studies with **1** demonstrated that this reverse orientation was feasible. Hydrogen-bonding to the backbone Gly216 of thrombin could be maintained in this parallel alignment of the backbone via Phe1 O and cHex-Ala3 N. The aromatic and the cyclohexyl residues bind at the P- and the D-pockets, respectively, while the hydroxy-thioimidazole is exposed to solvent.

Observations that are consistent with this model include: 1) replacement of the cyclohexyl by a phenyl group (known to bind well in the D-pocket) and replacement of the hydroxy-thioimidazole moiety by the simpler methyl ester moiety are tolerated (compound **5**, Table 1); 2) the analogs display sensitivity to the alkyl side chain length in both the amino and the guanidino terminated series, and a preference for the guanidino group over the amino group by about 40 fold.<sup>37</sup> Compound **6** having a 4-guanidinobutanoyl moiety was the most potent inhibitor in this set of analogs.<sup>1</sup> Further SAR studies led to **7** (BMS-183,507), where the *allo*-threonine side chain leads to a 16-fold increase in activity. This compound was modeled in greater detail and subsequently, its crystal structure was solved.

### Crystal Structure of BMS-183,507

#### Comparison with hirudin

The crystal structure of human  $\alpha$ -thrombin complexed with BMS-183,507 and hirugen indicates that the inhibitor does bind in the 'retro-binding' fashion.<sup>2</sup> The Phe1-*allo*-Thr2-Phe3 portion of the inhibitor binds like the first three N-terminal residues of hirudin. The inhibitor backbone aligns parallel to the main chain segment Ser214 to Glu217 of thrombin with Phe1 N forming a hydrogen bond to Ser214 O, Phe1 O to Gly216 N and Phe3 N to Gly216 O. The Phe1 and Phe3 side-chains bind in the P-pocket and D-pocket respectively, while the *allo*-Thr2 side chain lies above Gly219 (Fig. 5). The rms difference for BMS-183,507 versus Hir Ile1-Tyr3 is only 1.28 Å (using N, C $\alpha$ , C, C $\beta$ , O atoms). In contrast to the hirudin structure, BMS-183,507 possess two new interactions: the alkyl-guanidine moiety binds in the specificity pocket and the *allo*-Thr2 side chain hydroxyl oxygen forms a hydrogen bond to Gly219 N. Banner *et al.* has also reported that Gly219 N forms a hydrogen bond in another series of thrombin inhibitors.<sup>38</sup>

The crystal structure of BMS-183,507- $\alpha$ -thrombin offers additional insight into the structural requirements for

binding for these tripeptide inhibitors and for the N-terminal residues of hirudin. The loss in binding affinity of hirudin upon acylation of its N-terminus may be attributed to the loss of the positive charge, which leads to a loss in the hydrogen bond or a salt bridge between the  $\alpha$ -amino group of Ile1 (hirudin) and thrombin.<sup>20,21</sup> However, in the case of these retro-peptides, the loss of this N-terminal interaction is apparently compensated for by the new and stronger interaction between the protonated guanidine or amine terminus and Asp189 in the specificity pocket. In addition, modeling suggests that *N*-acetyl-hirudin would be sterically hindered. We found it difficult to place the N-terminus of the model of *N*-acetyl-hirudin into the active site because of the constraints imposed upon it by the rest of the hirudin molecule. However, the smaller tripeptides can avoid these steric contacts via a slight shift in their backbone as they possess greater rotational and translational freedom (Fig. 6).

#### Comparison with computational model

The modeled bound conformer for BMS-183,507 is similar to the crystal structure derived conformer (MC/EMII structure, Fig. 4, Table 2). However, the crystal structure indicates two new interactions that were not predicted from the model. The first interaction was the hydrogen bond formed by the *allo*-threonine residue. The hydroxyl of the *allo*-threonine forms a hydrogen bond to Gly219 N in the crystal structure (with  $\chi_1$  angle for the *allo*-threonine residue being  $-179^\circ$ ), but forms a hydrogen bond to Glu192 in the model. Although a hydrogen bond of the *allo*-threonine hydroxyl with the charged carboxyl group is stronger relative to the hydrogen bond with the neutral Gly219 N,<sup>39,40</sup> the Glu192 side chain in the BMS-183,507 crystal structure has moved away from the binding pocket into the solvent such that Gly219 N becomes the closest available hydrogen bonding partner for the hydroxyl of the *allo*-threonine.

The second new interaction was between the guanidine and the residues of the specificity pocket. The guanidine group of BMS-183,507 forms only one hydrogen bond to the Asp 189 carboxyl, N $\eta$ 1 to O $\delta$ 2, and the second nitrogen N $\eta$ 2 forms a hydrogen bond to the carbonyl O of Gly219. This interaction is probably facilitated by an intramolecular hydrogen bond between N $\epsilon$  of the guanidine with the carbonyl O of Phe1 of the inhibitor which results in a bent conformation for the alkyl chain. This guanidine interaction is different from the model and from the PPACK and NAPAP structures where the guanidine forms a pair of parallel bidentate hydrogen bonds (Type A, Fig. 7) with the carboxyl of Asp189 at the bottom of the specificity pocket.

#### Guanidine-carboxylate geometries

There are few high resolution crystallographic structures of enzyme-inhibitor-substrate complexes which involve arginine-carboxylate interactions. The availability of the different inhibitor-thrombin complexes allows us to compare the different ways in which a guanidine may interact with the same carboxyl (note that the structures

are at 3 Å resolution, or better). Some of the known geometries for a guanidine interacting with a carboxyl group are shown in Figure 7.

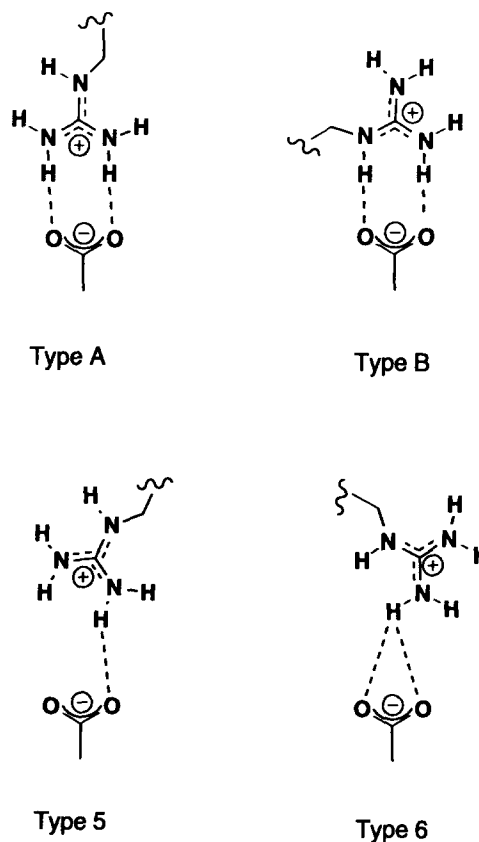


Figure 7. Hydrogen-bonding geometries for interacting arginine-carboxyls

The Type A geometry occurs frequently for the inhibitors of thrombin: the guanidine of PPACK and the fibrinopeptide substrate for thrombin, and the amidine of both benzamidine and NAPAP adopt this geometry in their interaction with the carboxyl of Asp189. This geometry occurs with high frequency in small molecule crystals (33%)<sup>41</sup>, but is rare in protein structures (less than 2%).<sup>42</sup> The Type B geometry has not been observed thus far for thrombin inhibitors, although it is frequently seen in protein structures (22%). Small molecule crystal structures reveal no clear preference for Type A over Type B, while packing in proteins appears to favor a Type B geometry. Quantum mechanical calculations on methylguanidinium-carboxylate indicate that the two forms are essentially identical energetically.<sup>43</sup> The guanidine of MD805 adopt a Type 6 geometry which is rare in proteins (less than 3%), while the guanidine of BMS-183,507 adopts a Type 5 geometry, which is the most frequently observed geometry in proteins (28%).<sup>42</sup>

The geometries of these guanidine-carboxylate interactions may be affected by the manner in which the alkyl or the aromatic functionalities having the guanidine/amidine moiety access the specificity pocket. The Type A geometry is seen in the inhibitors having the most direct approach of the alkyl/aryl-guanidine moiety into the specificity pocket, while the Type 5 and Type 6 geometries



are seen in the inhibitors which enter the specificity pocket obliquely. The guanidine/amidine of the inhibitors forms an additional hydrogen bond to the carbonyl oxygen of Gly219 via N $\eta$ 1 except for MD-805 which forms this hydrogen bond with Ne.

### Discussion

A series of retro-binding tripeptide inhibitors of thrombin was recently developed.<sup>1</sup> These inhibitors bind to the enzyme with their main chain in the opposite orientation to that described for thrombin substrates.<sup>2</sup> This reversed binding mode has also been reported for the structures of elastase complexed to TFA-Lys-Ala-*p*-trifluoromethylanilide,<sup>44</sup> TFA-Leu-Ala-*p*-trifluoromethylanilide<sup>45</sup> and TFA-Lys-Pro-*p*-isopropylanilide.<sup>46</sup> These elastase inhibitors form a parallel  $\beta$ -sheet with the protein, with the amino-terminus rather than the carboxy-terminus closest to the catalytic triad. Like the retro-peptides, the elastase inhibitors bind with their moieties occupying the substrate subsites. It is likely that this reversed binding mode will be seen in other serine proteases as more crystal structures of other inhibitor-serine proteases complexes become available.

Computational binding models for the inhibitors developed using a combination of DG/EM and MC/EM calculations led us to hypothesize that the inhibitors bind like the first three N-terminus residues of hirudin. This was subsequently verified by the solution of the crystal structure of the BMS-183,507-thrombin complex. The computational model of BMS-183,507-thrombin reproduced the crystal structure quite closely (Table 2, Fig. 4). It is encouraging that the minimum energy structure obtained from the docking studies only differed from the crystal structure in the interactions of the *allo*-threonine side chain and in the interactions of the arginine at the specificity pocket. Energy minimization optimized the geometries of the initial DGEOM complexes and allowed the identification of the structure closest to the experimental structure. Similar results have been observed in such hybrid approaches employing the DOCK suite of programs and minimization.<sup>47</sup>

These studies demonstrate the utility of modeling for understanding and guiding SAR studies of the tripeptide inhibitors. Modeling together with crystallographic studies which often reveals novel interactions that cannot be modeled accurately, provide powerful tools for structure-based drug design. Understanding the molecular basis for the inhibitor-enzyme interactions is a requirement for the rational design of new inhibitors. Together, the structures as shown in Figure 1, highlight the importance of maintaining three key binding interactions: 1) hydrophobic binding to the P- and D-pockets; 2) binding to the specificity pocket and forming salt bridges or ion pairs to Asp189; 3) forming a pair of hydrogen bonds to Gly216. The design of potent new small inhibitors of thrombin, especially those that do not form covalent intermediates with the enzyme, should include moieties that permit these interactions to occur. The hydrogen bond between the *allo*-threonine hydroxyl and Gly219 N is another inter-

action that can be incorporated into the design of new inhibitors.

### Acknowledgements

We would like to thank S. M. Seiler, D. G. Roberts and I. M. Michel for biological assays and for the identification of the lead compound. We thank S. D. Kimball for helpful discussions and for the substructure searching studies. We also thank C. Y. Chang and S. L. Ohringer for crystal preparation and L. Sieker and S. Sheriff for their assistance in the data collection, J. Lin, W.-C. Han and T. C. Wang for assistance in synthesis, and members of the Analytical Department at Bristol-Myers Squibb for providing analytical data. We also thank Dr W. Bode, Dr A. Tulinsky and Dr R. Huber for the coordinates of PPACK-human  $\alpha$ -thrombin and hirudin-human  $\alpha$ -thrombin.

### References and Notes

- Iwanowicz, E. I.; Lau, W. F.; Lin, J.; Roberts, D. G. M.; Seiler, S. M. *J. Med. Chem.* **1994**, *37*, 2111.
- Taberno, L.; Chang, C.-Y. Y.; Ohringer, S.; 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.
- Rydel, T. J.; Ravichandran, K. G.; Tulinsky, A.; Bode, W.; Huber, R.; Roitsch, C.; Fenton, II J. W. *Science* **1990**, *249*, 227.
- Rydel, T. J.; Tulinsky, A.; Bode, W.; Huber, R. *J. Mol. Biol.* **1991**, *221*, 583.
- Gutter, M. G.; Priestle, J. P.; Rahuel, J.; Grossenbacher, H.; Bode, W.; Hofsteenge, J.; Stone, S. R. *EMBO J.* **1990**, *9*, 2361.
- Fenton, II J. W. *Thrombin Ann. N.Y. Acad. Sci.* **1986**, *485*, 5.
- Fenton, II J. W.; Ofosu, F. A.; Moon, D. G.; Maraganore, J. M. *Blood Coagulation and Fibrinolysis* **1991**, *2*, 69.
- Berliner, L. T. *Thrombin Structure and Function*, Plenum Press; New York, 1992.
- Tapparelli, C.; Metternich, R.; Ehrhardt, C.; Cook, N. S. *TIPS* **1993**, *14*, 366.
- Bode, W.; Mayr, I.; Baumann, U.; Huber, R.; Stone, S. R.; Hofsteenge, J. *EMBO J.* **1989**, *8*, 3467.
- Bode, W.; Turk, D.; Karshikov, A. *Protein Sci.* **1992**, *1*, 426.
- Banner, D. W.; Hadvary, P. *J. Biol. Chem.* **1991**, *266*, 20085.
- Maryanoff, B. E.; Qiu, X.; Padmanabhan, K. P.; Tulinsky, A.; Almond, Jr H. R.; Andrade-Gordon, P.; Greco, M. N.; Kauffman, J. A.; Nicolaou, K. C.; Liu, A.; Brungs, P. H.; Fusetani, N. *Proc. Natl Acad. Sci. U.S.A.* **1993**, *90*, 8048.
- 'P-pocket' and 'D-pocket' as used by Banner and Hadvary.<sup>12</sup> The Tyr60A-Pro60B-Pro60C-Trp60D loop creates with Trp215, Leu99, and Ile174 a double hydrophobic pocket which they have divided into the P-pocket (Proximal to the active site serine) and a D-pocket (Distal to the active site serine). The P-pocket is also referred to as the S2 subsite.<sup>11,17</sup> The D-pocket is also referred to as the S9 subsite or the 'aryl-binding site'.<sup>11,17</sup> Numbering for thrombin residues is based on that for chymotrypsin.<sup>10</sup>
- Polgar, L. In: *Mechanisms of Protease Action*, pp. 87-122, Polgar, L., Ed.; CRC Press; Boca Raton, Florida, 1989.

16. Bode, W.; Huber, R. *Curr. Opin. Struct. Biol.* **1991**, *1*, 45.
17. Martin, P. D.; Robertson, W.; Turk, D.; Huber, R.; Bode, W.; Edwards, B. F. *J. Biol. Chem.* **1992**, *267*, 7911.
18. Goodman, M.; Chorev, M. In: *Perspectives in Peptide Chemistry*, pp. 283–294, Eberle, A.; Geiger, R.; Wieland, T., Eds; Karger, Basel, 1981.
19. Aguilar, C. F.; Thomas, P. J.; Moss, D. S.; Mills, A.; Palmer, R. A. *Biochim. Biophys. Acta* **1991**, *1118*, 6.
20. Stone, R. S.; Maraganore, J. M. In: *Thrombin Structure and Function*, pp. 219–256, Berliner, L. T., Ed.; Plenum Press; New York, 1992.
21. Wallace, A.; Dennis, S.; Hofsteenge, J.; Stone, S. R. *Biochemistry* **1989**, *28*, 10079.
22. Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comp. Chem.* **1990**, *11*, 440. MacroModel/BatchMin Version 3.5x was obtained from Professor C. Still, Columbia University, NY10027.
23. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, Jr S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
24. Guida, W. C.; Bohacek, R. S.; Erion, M. D. *J. Comp. Chem.* **1992**, *13*, 214.
25. Wipff, G.; Dearing, A.; Weiner, P. K.; Blaney, J. M.; Kollman, P. *J. Am. Chem. Soc.* **1983**, *105*, 997.
26. SYBYL Molecular Modeling System, Versions 5 and 6, TRIPOS Assoc., St Louis, MO, U.S.A.
27. Blaney, J. M.; Crippen, G. M.; Dearing, A.; Dixon, J. S. DGEOM, Program No. 590, Quantum Chemistry Program Exchange, Indiana University.
28. Blaney, J. M.; Dixon, J. S. In: *Reviews in Computational Chemistry*, pp. 299–335, Boyd, D. B., Ed.; VCH Publishers; New York, 1994.
29. Blaney, J. M.; Dixon, J. S. *Annu. Rep. Med. Chem.* **1991**, *20*, 281.
30. SAS, Version 5 available from SAS Institute, Inc., Cary, North Carolina.
31. Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379.
32. Fitzgerald, P. M. D. *J. Appl. Crystal.* **1988**, *21*, 273.
33. Sack, J. S. *J. Mol. Graphics.* **1988**, *6*, 244.
34. Brünger, A. T.; Kuriyan, J.; Karplus, M. *Science* **1987**, *235*, 458.
35. Iwanowicz, E. J.; Lin, J.; Roberts, D. G. M.; Michel, I. M.; Seiler, S. M. *Bioorg. Med. Chem. Lett.* **1992**, *12*, 1607.
36. Edwards, P. D.; Meyer, Jr E. F.; Vijayalakshmi, J.; Tuthill, P. A.; Andisik, D. A.; Gomes, B.; Strimpler, A. *J. Am. Chem. Soc.* **1992**, *114*, 1854.
37. Wigley, D. B.; Lyall, A.; Hart, K. W.; Holbrook, J. J. *Biochem. Biophys. Res. Commun.* **1987**, *149*, 927.
38. Banner, D.; Ackermann, J.; Gast, A.; Gubernator, K.; Hadváry, P.; Hilpert, K.; Labler, L.; Müller, K.; Schmid, G.; Tschopp, T.; van de Waterbeemd, H.; Wirz, B. In: *Perspectives in Medicinal Chemistry*, pp. 27–43, Testa, B.; Kyburz, E.; Fuhrer, W.; Giger, R., Eds; Verlag Helvetica Chimica Acta; Basel, VCH Basel, 1993.
39. Fersht, A. R. *Trends Biochem. Sci.* **1987**, *12*, 301.
40. Scholtz, J. M.; Qian, H.; Robbins, V. H.; Baldwin, R. L. *Biochemistry*, **1993**, *32*, 9668.
41. Salunke, D. M.; Vijayan, M. *Int. J. Peptide Protein Res.* **1981**, *18*, 348.
42. Singh, J.; Thornton, J. M.; Snarey, M.; Campbell, S. F. *FEB* **1987**, *224*, 161.
43. Deerfield, II D. W.; Nicholas, Jr H. B.; Hiskey, R. G.; Pedersen, L. G. *Proteins: Struct., Funct., Genet.* **1989**, *6*, 168.
44. Hughes, D. L.; Sieker, L. C.; Bieth, J.; Dimicoli, J.-L. *J. Mol. Biol.* **1982**, *162*, 645.
45. De la Sierra, I. L.; Papamichael, E.; Sakarellos, C.; Dimicoli, J. L.; Prange, T. *J. Mol. Recognit.* **1990**, *3*, 36.
46. Mattos, C.; Rasmussen, B.; Ding, X.; Petsko, G. A.; Ringe, D. *Nat. Struct. Biol.* **1994**, *1*, 55.
47. Meng, E. C.; Gschwend, D. A.; Blaney, J. M.; Kuntz, I. D. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 266.

(Received in U.S.A. 10 January 1995)