Structure-Based Understanding of Ligand Affinity Using Human Thrombin as a Model System

Vicki L. Nienaber,* Lawrence J. Mersinger, and Charles A. Kettner

Department of Chemical and Physical Sciences, The DuPont Merck Pharmaceutical Company, Experimental Station, Wilmington, Delaware 19880

Received September 11, 1995; Revised Manuscript Received April 11, 1996[⊗]

ABSTRACT: Kinetic study of a series of compounds containing the thrombin-directed peptide D-Phe-Pro-boroArg-OH had indicated that the structure of the N-terminal blocking group may be correlated with binding [Kettner, C., Mersinger, L., & Knabb, R. (1990) *J. Biol. Chem. 265*, 18289−18297]. In order to further study this phenomenon, a second series of compounds that contains a C-terminal methyl ester in place of the boronic acid was synthesized, binding measured, and the three-dimensional structure in complex with human thrombin determined by X-ray crystallography. Incubation of Ac-D-Phe-Pro-Arg-OMe, Boc-D-Phe-Pro-Arg-OMe, and H-D-Phe-Pro-Arg-OMe resulted in the formation of thrombin−product complexes within the crystal. K_i values for the corresponding products (free carboxylic acids) were 60 ± 12 μM, 7.8 ± 0.1 μM, 0.58 ± 0.02 μM, respectively, indicating that the nature of the N-terminal blocking group has a significant effect on affinity. Examination of the crystal structures indicated that the higher affinity of the H-D-Phe peptide is due to rearrangement of one residue comprising the S_3 site (Glu192) in order to maximize electrostatic interactions with the "NH_3^+−" of H-D-Phe. The relative affinity of Boc-D-Phe-Pro-Arg-OH is due to favorable hydrophobic interactions between thrombin and the bulky butyl group. However, this results in less favorable binding of Arg−P_1 in the oxyanion hole as shown by long hydrogen-bonding distances. This work gave rise to some general observations applicable to structure-based drug design: (1) altering the structure of an inhibitor at one site can affect binding at an unchanged distal site; (2) minor alteration of inhibitor structure can lead to small, but significant reorganization of neighboring protein structure; (3) these unexpected reorganizations can define alternate binding motifs.

The blood protein thrombin cleaves fibrinogen into insoluble fibrin in the final step of the coagulation cascade.[1] Thrombin also functions as a procoagulant by stimulating a number of other blood factors as well as platelet aggregation but will also act as an anticoagulant when complexed with thrombomodulin (Fenton, 1988; Stubbs & Bode, 1993; Davie *et al.*, 1991). Because thrombin plays a central role in the regulation of blood coagulation, it has been targeted in the treatment of various hemostatic disorders such as myocardial infarction, stroke, and pulmonary embolism (Fenton *et al.*, 1991).

Many synthetic thrombin inhibitors contain the D-Phe-Pro-Arg sequence. Bajusz *et al.* (1978) synthesized the tripeptide aldehyde, H-D-Phe-Pro-Arg-H and demonstrated that it was effective in blocking thrombin's cleavage of fibrinogen in the submicromolar range. The active site serine reversibly adds to the arginine aldehyde forming a tetrahedral complex similar to the tetrahedral intermediate expected for normal substrate hydrolysis. Kettner and Shaw (1979) synthesized the chloromethyl ketone, H-D-Phe-Pro-Arg-CH_2Cl (PPACK), which forms a complex with the active site serine similar to the peptide aldehyde and also irreversibly alkylates the active site histidine of thrombin. Finally, Kettner *et al.* (1990) synthesized the compound H-D-Phe-Pro-boroArg-OH,[2] which is one of the more potent synthetic thrombin inhibitors reported to date with a K_i of <3.6 pM. As with the aldehyde and chloromethyl ketone homologues, H-D-Phe-Pro-boroArg-OH forms a tetrahedral transition state-like complex with thrombin via a covalent bond between the boroarginine boron and the active site Ser195 hydroxyl.

Further studies of a X-D-Phe-Pro-boroarginine series indicated that the blocking group (X) present on D-phenylalanine effects the potency of the inhibitor (Kettner *et al.*, 1990). Specifically, these studies showed that H-D-Phe-Pro-boroArg-OH was the most effective inhibitor binding with a K_i of <3.6 pM. Boc-D-Phe-Pro-boroArg-OH binds with a K_i of 3.6 pM, and Ac-D-Phe-Pro-boroArg-OH has a K_i of 41 pM. A series consisting of the corresponding peptide methyl esters (Ac-D-Phe-Pro-Arg-OMe, Boc-D-Phe-Pro-Arg-OMe, and H-D-Phe-Pro-Arg-OMe) were incubated with thrombin, and the structures of the resulting complexes were determined. Each complex consisted of the enzyme bound to product. The dissociation constants of these product complexes correspond to the same trend observed for the boronic acid inhibitors. Crystal structures of complexes obtained by incubating the methyl ester compounds with human thrombin allowed correlation of these product−thrombin complex structures with binding affinity.

As reported by Bode *et al.* (1989), the overall fold of human α-thrombin is similar to that of the pancreatic serine proteases. Common features consist of two β-barrels, the

* Author to whom correspondence should be addressed. Current address: Department of Structural Biology, D46Y/AP10, Abbott Laboratories, 100 Abbott Park Road, Abbott Park, IL 60064. Phone: 847-935-0918. Fax: 847-937-2625. Email: nienabev@crow.pprd.abbott.com.

[⊗] Abstract published in *Advance ACS Abstracts,* June 15, 1996.

[1] The chymotrypsin numbering system as aligned by Bode *et al.* (1989) will be used.

[2] Abbreviations: boroArg-OH, arginine where the carboxyl group has been replaced by −$B(OH)_2$; Boc, *tert*-butyloxycarbonyl; Ac, acetyl.

S0006-2960(95)02164-7 CCC: $12.00

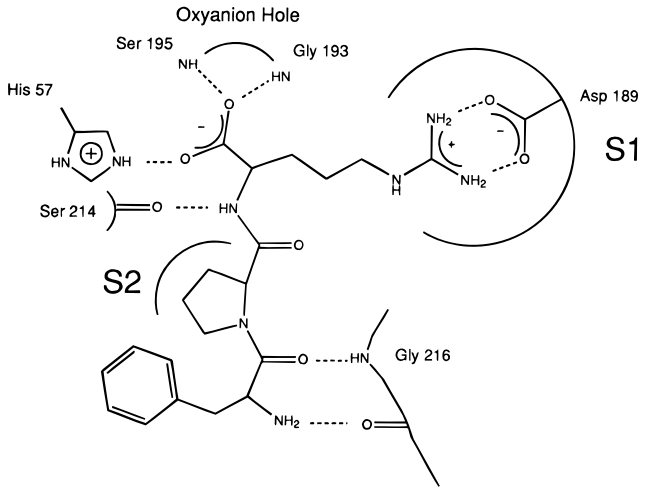


FIGURE 1: Schematic diagram of the typical hydrogen-bonding pattern observed between a peptide inhibitor and a serine protease. For trypsin-like proteases such as thrombin a negatively charged residue (Asp189) is found at the base of the S_1 pocket and is shown forming a salt bridge with Arg−P_1. Hydrogen bonds resulting in an antiparallel β-sheet interaction between the peptide and thrombin are also shown. The C-terminal carboxyl is expected to form two hydrogen bonds at the oxyanion hole (Ser195−N and Gly193−N) as well as a salt bridge with His57 of the catalytic triad.

catalytic triad (Ser195, His57, Asp102), and a substrate binding groove located between the two domains (Figure 1). At the substrate binding groove, a series of pockets[3] defines the cleavage specificity of the protease and invokes a substrate orientation ideal for enzymatic cleavage. Thrombin, a trypsin-like serine protease, prefers to cleave after positively charged lysine or arginine residues due to a negatively charged aspartic acid located at the base of the P_1 binding pocket (Figure 1). Thrombin also contains a unique apolar binding site at the P_2 position that is composed partially of a nine residue insertion loop at position 60. The side chains of Tyr60A and Trp60D contribute to the pocket along with those of His57 and Leu99. In general, a proline hydrophobic side chain is preferred at position P_2 because it tends to invoke the optimal geometry for formation of a series of hydrogen bonds (see Figure 1) which forms a stretch of antiparallel β-sheet between the substrate and one rim of the P_1 binding pocket (Kettner & Shaw, 1979; Bajusz *et al.*, 1978).

The structural basis for the preferred unnatural amino acid, D-phenylalanine, at position P_3[3] (Kettner & Shaw, 1979; Bajusz *et al.*, 1978) was also explained by Bode *et al.* (1989) in their presentation of the H-D-Phe-Pro-Arg-CH-thrombin crystal structure. The D-Phe side chain is in van der Waals contact with the P_2 proline ring and forms one side of the S_2 pocket. This D-phenylalanine−proline interaction has also been shown to exist in solution (Lim *et al.*, 1993), indicating entropic advantages in the binding of this inhibitor. Besides interactions with the P_2 side chain, the D-phenylalanine also interacts with Trp215 in an edge-to-face manner as described by Burley and Petsko (1988). This should also contribute toward a favorable binding energy between a P_3 D-phenyalanine residue and thrombin.

The interactions between the P_1, P_2, and P_3 side chains of the inhibitors presented here and their respective binding pockets on thrombin were nearly identical to the interactions first observed by Bode *et al.* (1989) for H-D-Phe-Pro-Arg-CH-thrombin. Thus, the results presented will primarily focus on the blocking group (S_3) binding site and interactions at the oxyanion hole.

MATERIALS AND METHODS

Chemical Synthesis. The methyl ester peptides were prepared by solution-phase peptide chemistry using the mixed anhydride procedure and coupling to ω-nitroarginine methyl ester. Final products were obtained by removing the nitro protecting group by catalytic hydrogenation.

Boc-D-Phe-Pro-Arg-OH was purchased from Bachem. H-D-Phe-Pro-Arg-OH was obtained by treating the Boc peptide with anhydrous trifluoroacetic acid. Ac-D-Phe-Pro-Arg-OH was obtained by treating H-D-Phe-Pro-Arg-OH with acetic anhydride in aqueous sodium bicarbonate. The reaction was acidified and applied to a 2.2 × 25 cm column of Vydac C4. The product eluted as a single peak by chromatography with a gradient from 0.1% TFA to 100% acetonitrile.

Kinetic Measurements. The interactions of the arginine peptides and peptide methyl esters with thrombin were determined at 25 °C in 0.10 M sodium phosphate buffer, pH 7.5, containing 0.20 M sodium chloride and 0.5% PEG 8000. Binding of the enzyme−product complexes was determined by measuring the K_is of Ac-D-Phe-Pro-Arg-OH, Boc-D-Phe-Pro-Arg-OH, and H-D-Phe-Pro-Arg-OH for thrombin using the chromogenic substrates S-2238 (H-D-Phe-Pip-Arg-pNA) and S-2366 (pyroGlu-Pro-Arg-pNA) and an enzyme level of 0.20 nM. Average K_m values (from at least five separate determinations) measured over substrate concentration ranges of 0.10−1.0 mM and 2.0−50 μM were 110 ± 20 μM and 1.1 ± 0.5 μM, respectively, for S-2366 and S-2238. Inhibition of the hydrolysis of 0.20 mM S-2238 was measured for at least three different concentrations of inhibitor over the range of 100−500 μM. Similarly, hydrolysis of both 0.20 and 0.40 mM S-2366 was measured for at least two different inhibitor concentrations over a range of 100−500 μM for the former two compounds and 10−50 μM for the latter. The equation for competitive inhibition was readily fit to the the data, yielding consistent values of K_i. This includes values determined over a range of inhibitor levels for separate substrates and also different substrate levels. Reported K_is are the average of at least five separate determinations.

Binding of the peptide methyl esters was determined by treating them as competitive inhibitors as described by Segal (1975). Conditions almost identical to those for the carboxylic acids were used. Hydrolyses of 0.20 mM substrates were linear for 30 min in both the presence and absence of the arginine methyl esters for both chromogenic substrates.

Crystallographic Studies. Human α-thrombin was purchased from Enzyme Research Laboratories (South Bend, IN) and used without further purification while the hirudin C-terminal peptide, residues 54−65, was synthesized in house by Rose Wilk of the DuPont Merck Biotechnology Peptide Synthesis Facility. The thrombin-hirudin complex

[3] The nomenclature used is that of Schecter and Burger (1967) where residues N-terminal to the scissile bond are designated as P_1, P_2, P_3, ..., P_n and residues C-terminal to the scissile bond are P_1', P_2', P_3', ..., P_n'. For an all L-amino acid peptide the corresponding binding sites on the protein are S_1, S_2, S_3, ..., S_n and S_1', S_2', S_3'...S_n'. For this paper, the D-Phe residue at position P_3 binds in the S_4 site, while the blocking group N-terminal to the D-Phe binds at the S_3 site. This convention was adopted to allow discussion of binding of both all L-, and D-containing peptides to thrombin.

Table 1: Comparison of the Binding of Peptide Methyl Esters and Peptide Boronic Acids to Thrombin[a]

R	K_i (μM) R-Arg-OH[b]	K_i (μM) R-boroArg-OH
Ac-D-Phe-Pro-	60 ± 12	41 ± 2
Boc-D-Phe-Pro-	7.8 ± 0.1	3.6 ± 0.6
H-D-Phe-Pro-	0.58 ± 0.02	<3.6

[a] All measurements were made at 25 °C and pH 7.5. Values of K_i(final) for boronic acid binding have been reported previously (Kettner *et al.*, 1990). [b] The interaction of the corresponding methyl esters with thrombin was determined by their evaluation as competitive inhibitors according to the method described by Segal (1975). By this method, the following kinetic binding constants were obtained: 9.6 ± 0.7, 2.6 ± 0.4, and 1.1 ± 0.1 μM.

Table 2: Crystallographic Data for Thrombin−Inhibitor Complexes[a]

	resolution (Å)	pos reflcns	ind reflcns	total obs	R_{merge}[b]	% I $>2\sigma$[c]
Boc-D-FPR	2.0	24 273	21 151	90 442	0.095	78.1
Boc-D-FPR-OMe	2.0	24 179	22 052	81 420	0.060	90.4
Ac-D-FPR-OMe	2.0	24 023	21 005	57 489	0.085	85.2
D-FPR-OMe	2.0	24 437	20 181	48 854	0.110	76.7

[a] For 1σ data. [b] $R_{merge} = \Sigma\Sigma(|F^2(i) - \langle F^2(h)\rangle|)/\Sigma\Sigma F^2(i)$. [c] Data used in the crystallographic refinement.

was crystallized in the space group C_2 (a = 70.67 Å, b = 72.9 Å, c = 73.0 Å, β = 100.4°) by the procedure of Skrzypczak-Jankun *et al.* (1991) with some modifications. To prepare thrombin−inhibitor complexes, thrombin crystals were soaked in a solution containing 2.0 mg of inhibitor in 250 μL of crystal stabilization buffer [0.58 M sodium phosphate, pH 7.2, 33% poly(ethylene glycol) MW 8000, 0.05 mM NaN_3] for at least 24 h prior to mounting crystals in capillaries. Inhibitors were initially solubilized in DMSO.

X-ray diffraction data for the thrombin-inhibitor complexes were collected on an R-axis Image Plate system mounted on a Rigaku RU-200 rotating anode generator operating at 100 mA and 50 kV. Twenty minute two degree oscillation frames were collected for a 120° coverage at two crystal orientations. The data were then reduced using software supplied by the manufacturer. Crystals were sufficiently stable in the X-ray beam so that only one crystal was needed to collect a complete data set on each complex. Table 2 lists the crystallographic data parameters for thrombin−product complex crystals.

Initial difference Fourier maps were computed using the structure factors and phases calculated from the refined coordinates of the ternary complex between Ac-D-Phe-Pro-boroArg-OH, human α-thrombin, and the hirudin peptide (Weber *et al.*, 1995) with Ac-D-Phe-Pro-boroArg-OH removed. Inhibitors were placed in the initial difference electron density maps using the interactive graphics package TURBO (Roussel & Cambillau, 1989) and the coordinates refined using the program package X-PLOR (Brunger, 1990) with the parameter sets parhcsdx.pro and tophcsdx.pro. Refinement of the complexes continued with alternate cycles of positional and individual *B*-factor energy minimization at 2.0 Å resolution. Ordered solvent molecules were located as positive peaks which were at least 4σ above noise in difference electron density maps during the course of the refinement. The final refinement statistics are listed in Table 3. All structures are well refined as indicated by the crystallographic *R*-factors and acceptable geometric rms deviations.

Table 3: Refinement Statistics

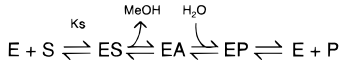
	Boc-D-FPR	Boc-D-FPR-OMe	Ac-D-FPR-OMe	D-FPR-OMe
R-factor[a]	0.191	0.194	0.181	0.223
resolution (Å)	8.0−2.0	8.0−2.0	8.0−2.0	8.0−2.0
rms bond length (Å)	0.017	0.016	0.017	0.021
rms bond angle (deg)	2.85	2.69	2.61	2.82
av B (Å^2)				
substrate	35.54	32.99	26.68	24.21
protein	27.68	28.99	29.45	32.15
solvent	38.27	48.37	51.92	53
no. of solvent atoms	133	133	148	116

[a] Crystallographic *R*-factor for data greater than 2σ.

Electrostatic calculations were performed at neutral pH using the UHBD algorithm as implemented in QUANTA (Davis *et al.*, 1991). This program utilizes the finite difference method to solve the linearized Poisson−Boltzmann equation. Partial atomic charges were taken from the CHARMM force field. For these calculations, the protein was assigned a dielectric constant of 2 while the solvent dielectric constant was taken to be 78. The grid size for the calculations was 1 Å, the box size was 100 Å, the ionic strength was 10 mM, and the temperature was 300 K.

RESULTS

General Aspects of Inhibitor Binding. The peptide methyl esters presented in this study were introduced to the thrombin crystals (E) as substrates (S). The methyl esters were hydrolyzed to yield an acyl-enzyme (EA) which is further hydrolyzed to an enzyme−product complex (EP). All structures were analyzed as EP complexes.



The methoxy group was not visualized in any of the electron density maps, indicating that the group had been enzymatically removed within the crystal lattice. In the EP complex, the carboxylate oxygen and His−Nϵ_2 are within hydrogen-bonding distance and the carbonyl oxygen in the oxyanion hole where hydrogen bonds are formed with Gly193−N and Ser195−N (see Table 4 and Figure 1). These interactions were observed for all esters in this series. Discrete differences between product complexes are discussed individually and summarized in Table 4.

Because the crystal structures reported are of EP complexes, the binding constants for both the substrates (see legend to Table 1) and product inhibitors (K_i) were determined (Table 1). For both series, the binding trend was similar to that observed with the boroarginine compounds.

The formation of a stable EP complex is not unique. The hydrogen-bonding pattern between the X-D-Phe-Pro-Arg carboxylates and thrombin is identical to that observed for the carboxylate of the noncovalent complex of residues 7−16 of fibrinopeptide A and bovine thrombin as determined by Martin *et al.* (1992). In addition, similar carboxylate−protein interactions have been observed at the active site of the glutamic acid specific serine protease (Nienaber *et al.*, 1993) from *Streptomyces griseus* bacteria as well as *S. griseus* protease A (James *et al.*, 1980).

General interactions between the P_1, P_2, and P_3 side chains of the product and human thrombin were as predicted on the basis of the crystal structure of H-D-Phe-Pro-Arg-CH-thrombin (Bode *et al.*, 1989). In all complexes, the positively

Table 4: Intermolecular Hydrogen Bond Distances in Thrombin−Inhibitor Complexes (Å)

				Boc-D-FPR	Boc-D-FPR-OMe	Ac-D-FPR-OMe	D-FPR-OMe	D-FPRCK[a]	Ac-D-FPboroR-OH[b]
R-P_1	N	S214	O	3.4	3.3	3.2	3.1	2.9	3.2
	OT1	S195	N	3.8	3.5	3.1	3.1	3.3	3.2
	OT1	G193	N	3.4	3.4	3.1	2.9	3.2	2.9
	OT2	H57	Nϵ_2	2.9	2.8	2.8	2.8	np[c]	2.6
	NH1	D189	Oδ_1	3.0	3.1	3.1	3.3	2.8	2.8
	NH2	D189	Oδ_2	2.6	2.6	2.7	2.8	2.6	2.6
	NH2	G219	O	2.8	2.8	2.9	3.1	2.8	2.6
D-F-P_3	O	G216	N	3.4	3.3	3.2	3.0	2.7	2.9
	N	G216	O	3.1	3.1	3.1	2.7	3.1	3.2

[a] Bode *et al.* (1989). [b] Weber *et al.* (1995). [c] Not present.

charged P_1 arginine forms a salt bridge with Asp189 at the base of the S_1 pocket; P_2 proline binds in the hydrophobic S_2 pocket, and P_3 D-phenylalanine caps the S_2 pocket while interacting with Trp215 of thrombin. In addition, all peptides form three main chain−main chain hydrogen bonds with residues Ser214 and Gly216 as illustrated in Figure 1. All intermolecular hydrogen bond distances between the three peptides and thrombin are summarized in Table 4.

H-D-Phe-Pro-Arg. The structure of H-D-Phe-Pro-Arg-thrombin was similar to that of H-D-Phe-Pro-Arg-CH-thrombin (Bode *et al*., 1989) with the major difference being alkylation of the active site histidine by the latter. In Figure 2A, a $2F_o - F_c$ map of D-Phe-Pro-Arg bound to human thrombin at 2.0 Å resolution clearly defines all inhibitor and protein atoms at the active site. Also, the absence of electron density between Ser195−OH and Arg−P_1 confirmed the presence of a noncovalent EP complex. The structure of D-Phe-Pro-Arg thrombin, as well as the hydrogen-bonding pattern, illustrated schematically in Figure 1, is shown in Figure 2B. Here, the hydrogen bonds which result in an antiparallel β-sheet between the inhibitor and one rim of the P_1 pocket are shown along with the salt bridge interaction between Arg-P_1 and Asp189 at the base of the S_1 pocket. In comparing the chloromethyl ketone structure with the hydrolyzed methyl ester, a single oxygen occupies the oxyanion hole (Ser195−N and Gly193−N) for both. The oxyanion hole has been proposed to play a major role in negative charge stabilization during the catalytic mechanism of serine proteases (Henderson, 1970). In the D-Phe-Pro-Arg complex, the unique carboxylate oxygen is accepting a hydrogen bond from His57−Nϵ_2 as well as from an ordered solvent molecule. This ordered solvent is in turn accepting a hydrogen bond from Lys60F−Nζ.

Ac-D-Phe-Pro-Arg. Figure 3A illustrates that the hydrogen-bonding pattern between Ac-D-Phe-Pro-Arg and thrombin is nearly identical to that observed with H-D-Phe-Pro-Arg thrombin. Consequently, the two compounds bind in a similar manner as depicted in Figure 3B.

In Figure 3C, the structure of Ac-D-Phe-Pro-Arg thrombin and Ac-D-Phe-Pro-boroArg-thrombin are overlaid to better compare their respective bound structures. Unlike the free carboxyl compound, Ac-D-Phe-Pro-boroArg, binds to thrombin via an additional covalent bond between Ser195−OH and the boron atom of boroArg−P_1. Because of this, the Cα position of boroArg−P_1 is shifted in towards the protein by about 0.5 Å relative to Arg−P_1. This change in binding at the S_1 site seems to have resulted in a different binding mode for the acetyl group of Ac-D-Phe-Pro-boroArg relative to Ac-D-Phe-Pro-Arg. For the noncovalent compound, the acetyl group is involved in hydrophobic interactions with thrombin Glu217−Cα, −Cβ, and −Cγ while for the boronic acid structure, the acetyl group is solvent exposed. Hence, in the boronic acid derivative, this hydrophobic interaction is compromised relative to the noncovalent free carboxyl compound. This is because the free carboxylate compound is apparently able to simultaneously maintain hydrogen-bonding interactions at the oxyanion hole and hydrophobic interactions at the acetyl site while the covalently linked boronic acid cannot. This is the first example presented where the binding at one site can affect inhibitor binding at a different site of the protein. The second example will be discussed below. This subtle change in inhibitor structure also alters the conformation of Glu192 of the thrombin molecule. In the boronic acid compound where the acetyl group is pointing out, this side chain is oriented in what will be referred to as the "up" position while in the free carboxyl compound, this side chain is pointing "down" (see Figure 3C).

Boc-D-Phe-Pro-Arg. The crystal structure of the Boc-D-Phe-Pro-Arg−thrombin complex showed that this compound binds in a manner different from the acetyl and free amine compounds. Because of this unusual result, the Boc compound was soaked into thrombin crystals in both the methyl ester (Boc-D-Phe-Pro-Arg-OMe) and free acid (Boc-D-Phe-Pro-Arg-OH) form. Examination of these two crystal structures allowed confirmation of this result and allowed for an internal control in comparing the hydrogen-bonding distances for all four structures (see Table 4). The two Boc compounds bound nearly identically with the differences in hydrogen-bonding distances between each of these compounds and thrombin being within error (±0.2 Å) for a 2.0 Å resolution structure. Hydrogen-bonding interactions (see Figure 4A) which are maintained are the β-sheet interactions between the peptide and protein and the salt bridge between Arg−P_1 and Asp189. Interactions between one oxygen of the C-terminal carboxyl and His57−Nϵ_2 and the ordered solvent molecule are also maintained, although the hydrogen bond between this solvent molecule and Lys60F−Nζ is broken. This results in Lys60F−Nζ rotating back such that it is close to Cys44−O (3.7 Å) and is donating a hydrogen bond to His43−O (2.9 Å). In the acetyl and free amine structures, the distance between Lys60F−Nζ and His43−O (3.9 Å) or Cys44−O (4.5 Å) is longer. This result indicates that changing the inhibitor structure at one site can affect the protein structure at a distal site nearly 15 Å away.

The second major difference between the binding of the Boc compound versus the acetyl or free amine is in the hydrogen bonding of the second oxygen of the C-terminal carboxyl. This oxygen is normally bound in the oxyanion hole (Ser195−N and Gly193−N). However, in both Boc structures, the hydrogen bond distances at this site are longer than those typically observed (see Table 4). This change in

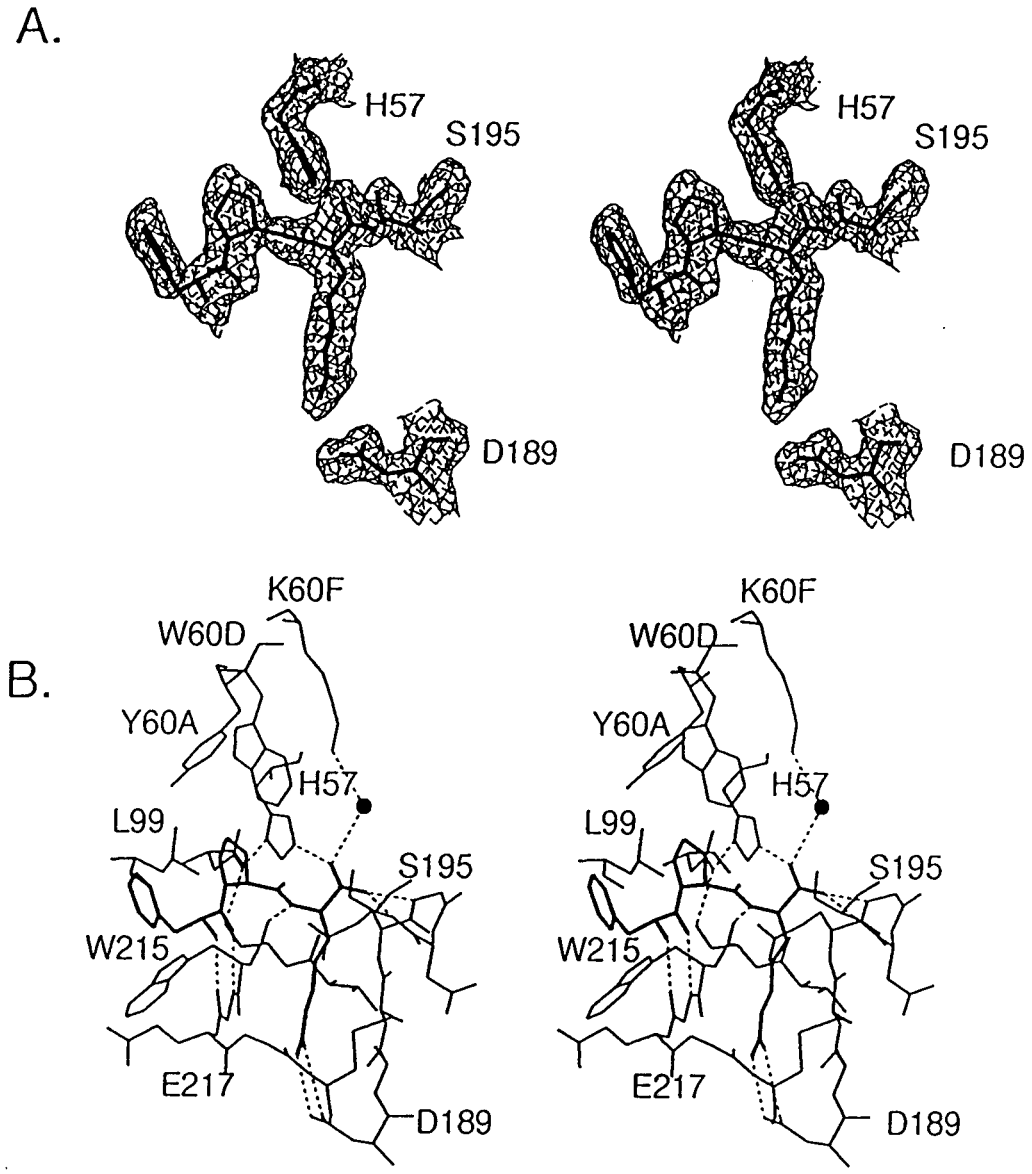


FIGURE 2: Stereo representation of the active site of D-Phe-Pro-Arg-thrombin. (A) $2F_o - F_c$ map contoured at 1σ around the inhibitor, Asp189 of the P_1 pocket, and Ser195 and His57 of the catalytic triad. (B) Hydrogen bonding pattern between the peptide and thrombin. The inhibitor is shown in the thicker lines; hydrogen bonds are represented by dashed lines. The sphere represents an ordered solvent molecule.

hydrogen-bonding pattern results from about a 0.5 Å shift in the C and Cα positions of the Boc peptide relative to the acetyl and free amine compounds. In the same manner as the acetyl compound (see Figure 4B), the Boc group of Boc-D-Phe-Pro-Arg is interacting with Glu217−Cα, −Cβ, and −Cγ although the interactions are more extensive because of the larger size of Boc. Thus, it is possible that the C-terminal carboxyl is removed from the oxyanion hole to maximize the hydrophobic contact between Boc and Glu217. Consequently, this result shows that changing an inhibitor structure at one site can alter binding to the protein at a different site even though the structure of the inhibitor at that site is unchanged.

The shift in inhibitor binding observed with the Boc compound is perhaps more dramatic than that observed in comparing the Ac-D-Phe-Pro-Arg thrombin and Ac-D-Phe-Pro-boroArg-thrombin structures discussed above. The difference between the acetyl structures results from covalent bond formation, while results with the Boc compounds suggest that hydrophobic interactions are maximized at the expense of hydrogen bonding.

The final difference between the Boc structure and that of the acetyl and free amine is in the orientation of Glu192. When Boc-D-Phe-Pro-Arg is bound to thrombin, this side chain is oriented in the "up" position while it is oriented in the "down" position when Ac-D-Phe-Pro-Arg and D-Phe-Pro-Arg are bound. To better understand the orientation of this side chain and its effect on inhibitor binding, electrostatic calculations were performed for both conformations.

Electrostatics. Overall, results of the electrostatic calculations were similar to those observed by Bode *et al.* (1992). As shown in Figure 5A, thrombin has two characteristic positively charged lobes located at the fibrinogen recognition site and the proposed heparin binding site. Thrombin also has a weak negatively charged field located through the center of the molecule. Calculations were performed on both the Boc−peptide complex and that of the free amine in the absence of both the hirudin peptide and the inhibitor. In Figure 5B,C, the effect of the orientation of Glu192 on the potential field at the blocking group $(S_3)^3$ site is shown. As expected, a net negative potential is predicted at this site when the glutamic acid is in the down position while the

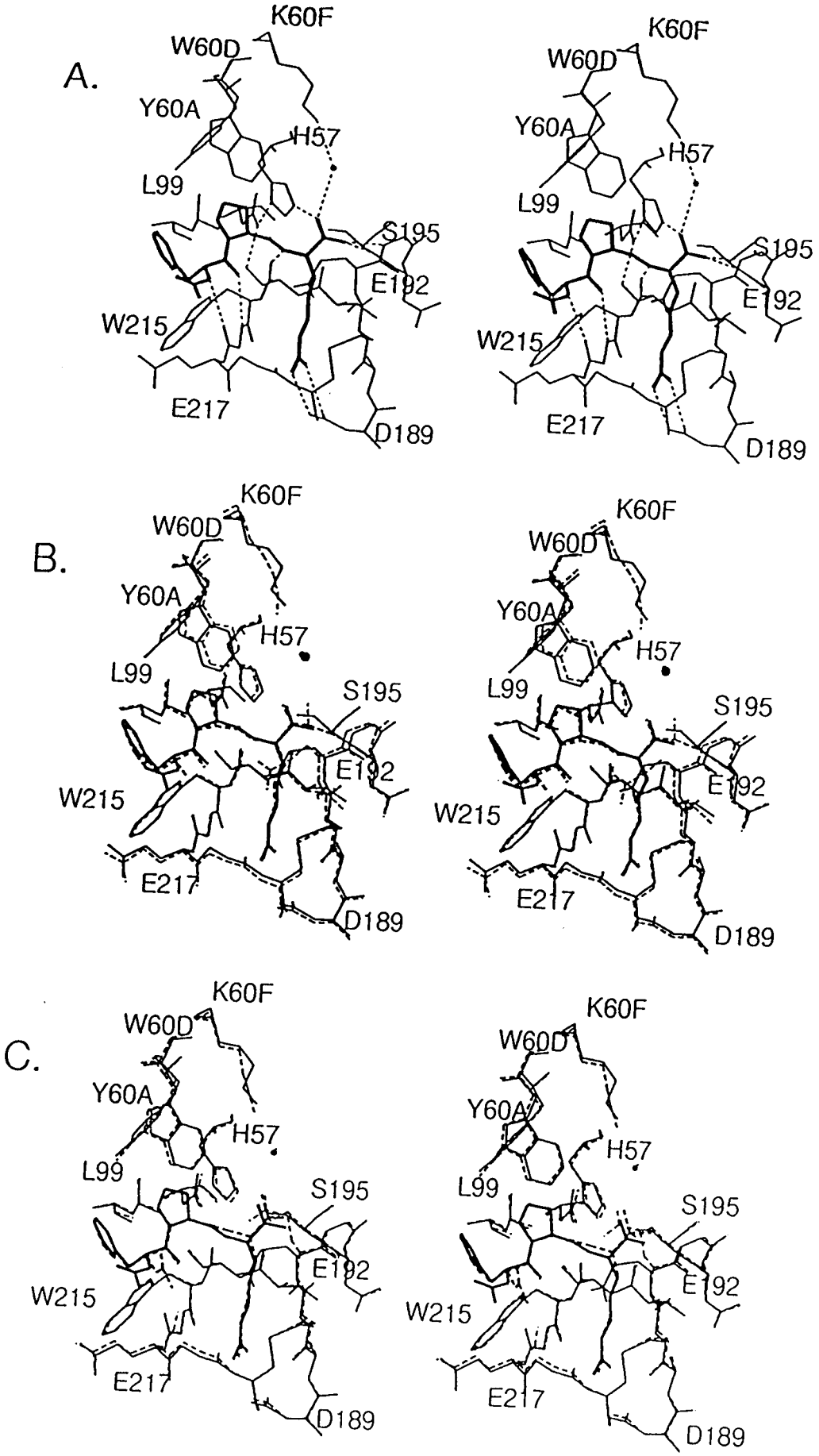


FIGURE 3: (A) Stereo representation of the hydrogen-bonding pattern between Ac-D-Phe-Pro-Arg (thick lines) and thrombin (thin lines). Hydrogen bonds are represented by dashed lines. (B) Overlay of D-Phe-Pro-Arg (dashed lines) and Ac-D-Phe-Pro-Arg (solid lines) bound to thrombin. Inhibitors are represented by the thick lines. (C) Overlay of Ac-D-Phe-Pro-Arg and Ac-D-Phe-Pro-boroArg. The boronic acid derivative structure is represented by dashed lines while the free carboxylate is shown in solid lines. Spheres represent ordered solvent molecules.

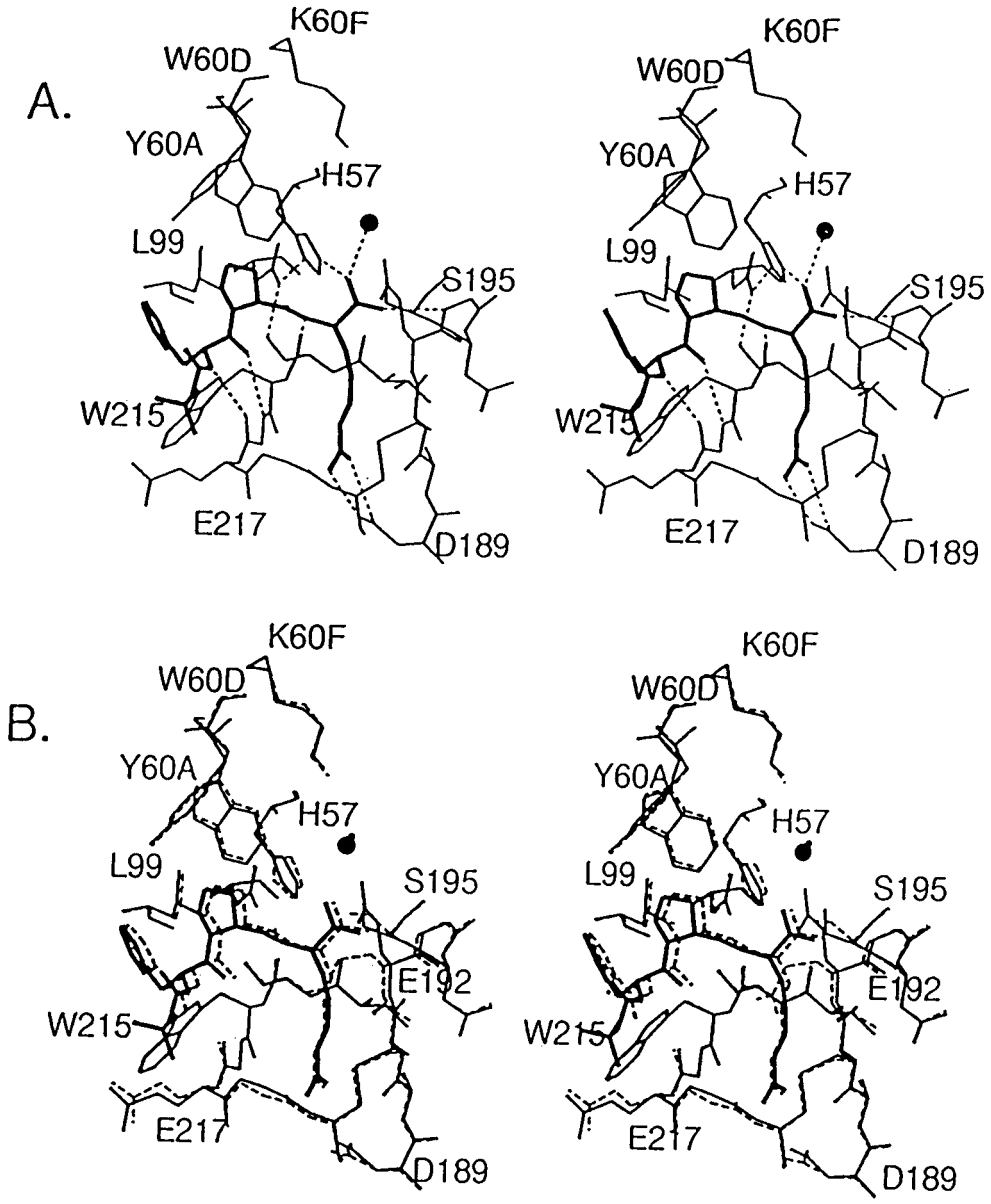


FIGURE 4: (A) Stereo representation of the hydrogen-bonding pattern between Boc-D-Phe-Pro-Arg (thick lines) and thrombin (thin lines). (B) Overlay of Boc-D-Phe-Pro-Arg (solid lines) and Ac-D-Phe-Pro-Arg (dashed lines). Spheres represent ordered solvent molecules.

site is left relatively neutral when this side chain is in the up position. Thus, the S_3 binding site of thrombin may prefer to bind neutral, hydrophobic residues when Glu192 is in the up position but may prefer positively charged side chains when this residue is oriented in the down position.

DISCUSSION

Crystal structures for a series of compounds complexed with human thrombin were determined to explore correlations between binding affinity and protein−inhibitor structure. Protein−inhibitor interactions which may contribute to binding energy include covalent bonds, hydrogen bonds, salt bridges, and hydrophobic interactions. Thus, in order to correlate inhibitor−protein structure and binding affinity, the interactions listed above were quantitated for each complex studied and the results correlated with acceptable energetic values.

Experimentally, the binding energy of a buried salt bridge has been determined to be 2.9 kcal mol^{-1} (Fersht, 1972) while that of a solvent-exposed salt bridge contributes less energy at about 1−2 kcal mol^{-1} (Perutz, 1970). The energetic contribution of a neutral−neutral hydrogen bond and a charged neutral hydrogen bond was measured at 0.5−1.5 and about 3 kcal mol^{-1}, respectively (Fersht, 1985). The relationship between change in free energy associated with removing a hydrophobic group from water and the surface area of that group has been shown to be linear (Reynolds *et al.*, 1974; Nozaki & Tanford, 1971; Chothia, 1974). This allows calculation of a proportionality constant detailing the free energy gained per hydrophobic surface buried. A rough estimate of this value for proteins has been proposed to be 20 cal $Å^{-2}$ mol^{-1} by Richards (1977).

Recently, a second statistical approach has been developed by Böhm (1994), who constructed an empirical scoring function to estimate the binding constant between small molecules and proteins based on the crystal structures of several protein−ligand complexes. His analysis includes the interactions mentioned above in addition to an entropic term resulting from the number of degrees of freedom lost upon ligand binding to the protein. His results indicate that the energy gained for an ideal neutral hydrogen bond was 1.1 kcal mol^{-1}, an ionic interaction was 2.0 kcal mol^{-1},

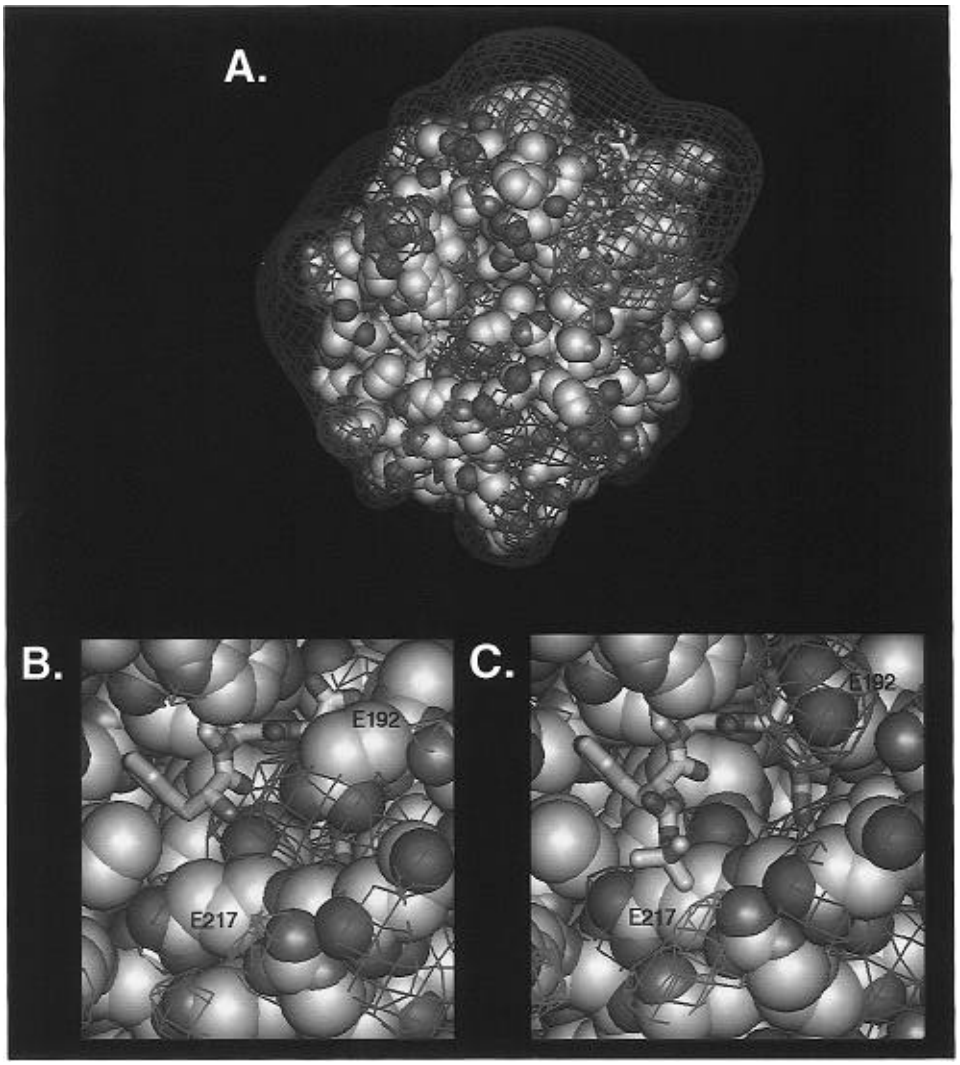


FIGURE 5: (A) Electrostatic potential map of human thrombin calculated in the absence of the hirudin peptide and the inhibitor. Thrombin is represented by a space-filling model, while the hirudin peptide and D-Phe-Pro-Arg inhibitor are shown as sticks. Blue represents positive potential contoured at 1 kT, and red represents negative potential contoured at −1 kT. (B) Structure of D-Phe-Pro-Arg-thrombin showing Glu192 in the "down" position overlaid on the electrostatic potential map. This shows a negatively charged S_3 blocking group site. (C) Structure of Boc-D-Phe-Pro-Arg-thrombin showing Glu192 in the "up" position overlaid on the electrostatic potential map. This shows a relatively neutral S_3 blocking group site.

hydrophobic contact contributed 40 cal $Å^{-2}$ mol^{-1}, and freezing out one rotatable bond decreases binding by 0.3 kcal mol^{-1}. Thus, Böhm's numbers for interactions between polar groups agree with those determined experimentally, while the number for nonpolar group interactions is twice that stated by Richards (1977). Both sets of numbers have been applied to the binding analysis of the compounds presented here.

In this study, the number of hydrogen bonds and salt bridges was the first parameter compared for each complex. With the exception of hydrogen bonding at the oxyanion hole (Ser195−N and Gly193−N), all hydrogen bonds and ionic interactions were nearly identical for the three compounds. However, for the Boc-peptide, longer hydrogen-bonding distances between the carbonyl oxygen and the oxyanion hole were observed. This is probably due to constraints arising from maximization of hydrophobic contact between the Boc protecting group and thrombin. However, in addition to binding at the oxyanion hole, the carboxylate of the cleaved methyl ester compounds is also accepting a hydrogen bond from a protonated His57−Nϵ_2 (Kossiakoff & Spencer, 1981). This histidine is probably the primary source for stabilization of the negatively charged C-terminus for this series of substrates. Thus, the loss of a hydrogen bond at the oxyanion hole may not significantly lower the binding energy of the Boc-peptide.

Table 5: Summary of Structural Interactions

	Ac	Boc	FRAM
no. of H-bonds (intermolecular)	9	8	9
hydrophobic buried ($Å^2$)			
total	632	724	560
inhibitor	508	567	453
protein	124	157	107

In addition to counting the number of hydrogen bonds, the hydrophobic surface buried upon inhibitor binding was also determined using the algorithm of Lee and Richards (1971) as implemented in X-PLOR (Brunger, 1990). As expected by the size of the N-terminal groups, this analysis indicated that the Boc-peptide buries the most hydrophobic area, the Ac-peptide the second while the free amine peptide buries the least (see Table 5). These results suggest that the Boc-peptide should bind most tightly, the acetyl second, and the free amine most weakly. This was also the result obtained when the analysis of Böhm (1994) was applied to this system. Contrary to these predictions, the free amine-

peptide binds most tightly. To further study this phenomenon, the protein structure at the blocking group (S_3) binding site was examined more closely.

As shown in Figure 4C, the most noticeable change in protein structure upon the binding of different inhibitors is the location of the side chain of Glu192. For the acetyl and Boc structures, the side chain is oriented in the up position, while for the free amine structure, the side chain is pointing down. This negatively charged amino acid residue is not hydrogen bonded to a positively charged residue in either position so it is expected to have a net negative charge. Electrostatic potential calculations support this negative potential and indicate that there is a net negative charge located at the S_3 site of thrombin when Glu192 is in the down position (see Figure 5C). Thus, it appears that this negative potential could balance the positive potential associated with the N-terminal free amine, resulting in the enhanced binding affinity of the free amine-peptide.

Long-range ion-pair interactions are not usually taken into account in structure-binding affinity correlations. However, it is important to note that they can contribute significant binding energy in contrast to long-range hydrophobic or polar hydrogen-bonding interactions. This is because for long-range ion-pair interactions, the energy varies as a function of $1/r$, where r is the distance between the two charged species. For noncharged, polar interactions which vary as $1/r^3$ and London forces which vary as $1/r^6$, the energy drops off more rapidly as a function of distance [reviewed by Burley and Petsko (1988)]. Furthermore, in their analysis of ion pairs in proteins, Barlow and Thorton (1983) found that the average distance between oppositely charged groups within the core of a protein is approximately 4.0 Å. Thus, long-range electrostatic interactions are found in the interior of proteins more often than those of a direct salt bridge. In this study, the distance between the free amine group and the carboxylate of Glu192 is 4.4 Å, which is within the range of the ion-pair interactions discussed by Barlow and Thorton (1983). In addition, if the energetic contribution of a salt bridge is corrected for the increase in distance (from about 3.3 to 4.4 Å), the relative binding energy for the free amine compound may be estimated. However, because this ion pair is partially solvent exposed, it is expected that the contribution to binding energy would be less than that if the ion pair was buried. In the free amine complex structure, the two charged species also appear to be connected via a hydrogen-bonding network of ordered solvent molecules. These solvent molecules may also contribute to binding affinity although they were not among the strongest solvent peaks in the original difference Fourier, suggesting that they may not be the sole contributor. Thus, both long-range electrostatic and water-mediated hydrogen-bonding forces may contribute to tighter binding of the free amine peptide.

The significance of the net negative potential arising from Glu192 has been suggested by Le Bonniec and Esmon (1991), who mutated Glu192 to a glutamine and looked at the effect of this mutation on thrombin's ability to cleave protein C. For thrombin to efficiently cleave protein C, it first must bind to the cofactor, thrombomodulin. However, Le Bonniec and Esmon (1991) found that Gln192−thrombin cleaves protein C with the same efficiency as the thrombin−thrombomodulin complex. This observation may be explained by the fact that the activation peptide of protein C has a negative charge at the P_3 (blocking group binding site for the D-Phe containing peptides[3]) site. Hence, in the absence of thrombomodulin, the negative potential of Glu192 could interfere with the cleavage of protein C while this effect is blocked in the thrombin−thrombomodulin complex. The results of Le Bonniec and Esmon (1991) together with the results presented here suggest that Glu192 of thrombin may play an important role in ligand binding and in regulation of thrombin activity. Furthermore, Rezaie and Esmon (1995) have recently shown that residue 192 may play a similar regulatory role in factor Xa specificity.

In summary, structural interactions between the protein and ligand were initially analyzed by conventional means, which consisted of counting the number of hydrogen bonds and calculating the buried hydrophobic surface area upon ligand binding. These contributions led to a predicted order of binding, Boc > acetyl > free amine, which is governed by the differences in the blocking group structure and which does not correlate with experimental data (see Table 1 and Table 5). For both series, the Boc compound binds more tightly than the acetyl which could be due to a larger hydrophobic surface buried for the Boc compound upon binding to thrombin, as indicated by crystal structures of the product complexes. The free amine compound was predicted to bind the most weakly because it lacks a blocking group although experimentally it was found to bind with affinity at least equivalent to the Boc compound for the methyl esters and boronic acids and was a significantly tighter binder in the case of the product complexes. Further examination of the structure indicated that binding of the free amine is potentiated by long-range electrostatic interactions with Glu192 of the protein which may be mediated by solvent molecules. Thus, the Boc and free amine compounds bind with similar affinities but by different mechanisms. This result also indicates that although long-range electrostatic interactions are complex to model theoretically, especially in the case of a solvent-exposed interactions, it is important to incorporate them into any qualitative correlation between binding affinity and structure.

This study also reminds us that crystal structures are time-averaged pictures of dynamic systems. This is an important point especially in the field of structure-based drug design where investigations often become very focused on the exact dimensions of microscopic protein−ligand interactions. It was shown that changing inhibitor structure at one site can alter binding at a second site (Boc structure versus the free amide and acetyl structures) in addition to altering the hydrogen-bonding pattern of a side chain on the protein up to 15 Å away (Lys60F−Nζ). This study also demonstrates how subtle movement of a single amino acid side chain can greatly alter the specificity of a binding site. Movement of Glu192 changes the specificity of the blocking group binding site from hydrophobic to cationic groups. Hence, in any predictive algorithm for ligand binding, these dynamic protein changes should be taken into account in addition to quantitating interactions between the protein and ligand. Thus, the resolution and accuracy of protein structures are only partial determinants for the prediction of binding of unknown inhibitors.

ACKNOWLEDGMENT

The authors thank Drs. Stephen Betz and Richard Alexander for critical evaluation of the manuscript as well as Drs. Pieter Stouten and Nicholas Hodge for helpful discussions. We also thank Frank Lewandowski for growing the thrombin

crystals used in this study and the DuPont Merck postdoctoral program for financial support of V. Nienaber.

-

REFERENCES

Bajusz, S., Barabas, E., Szell, E., & Bagdy, D. (1978) *Int. J. Pept. Protein Res. 12*, 217−221.

Barlow, D. J., & Thorton, J. M. (1983) *J. Mol. Biol. 168*, 867−885.

Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S., & Hofsteenge, J. (1989) *EMBO J. 8*, 3467−3475.

Bode, W., Turk, D., & Karshikov, A. (1992) *Protein Sci. 1*, 426−471.

Böhm, H. J. (1994) *J. Comput.-Aided Mol. Des. 6*, 243−256.

Brunger, A. T. (1990) *X-PLOR* (*version 2.1*) *Manual*, Yale University, New Haven, CT.

Burley S. K., & Petsko, G. A. (1988) *Adv. Protein Chem. 39*, 125−189.

Chothia, C. (1974) *Nature 248*, 338−39.

Davie, E. W., Fujikawa, K., & Kisiel, W. (1991) *Biochemistry 30*, 10363−10370.

Davis, M. E., Madura, J. D., Luty, B. A., & McCammon, J. A. (1991) *Comput. Phys. Commun. 62*, 187−197.

Esmon, C. T. (1989) *J. Biol. Chem. 264*, 4743−4746.

Fenton, J. W., II (1988) *Semin. Thromb. Hemostasis 14*, 234−240.

Fenton, J. W., II., Ofosu, F. A., Moon, D. G., & Maraganore, J. M. (1991) *Blood Coagulation Fibrionolysis 2*, 69−75.

Fersht, A. R. (1972) *J. Mol. Biol. 64*, 497−509.

Fersht, A. R., Shi, J., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M, & Winter, G. (1985) *Nature 314*, 235−238.

Henderson, R. (1970) *J. Mol. Biol. 54*, 341−354.

James, M. N. G., Sielecki, A. R., Brayer, G. D., & Delbaere, L. T. J. (1980) *J. Mol. Biol. 144*, 43−88.

Kettner, C. A., & Shaw, E. (1979) *Thromb. Res. 14*, 969−973.

Kettner, C., Mersinger, L., & Knabb, R. (1990) *J. Biol. Chem. 265*, 18289−18297.

Kossiakoff, A. A., & Spencer, S. A. (1981) *Biochemistry 20*, 6462−6474.

Le Bonniec, B. F., & Esmon, C. T. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 7371−7375.

Lee, B. and Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.

Lim, M. S. L., Johnston, E. R., & Kettner, C. A. (1993) *J. Med. Chem. 36*, 1831−1838.

Martin, P. D., Robertson, W., Turk, D., Huber, R., Bode, W., & Edwards, B. F. P. (1992) *J. Biol. Chem. 267*, 7911−7920.

Nienaber, V. L., Breddam, K., & Birktoft, J. J. (1993) *Biochemistry 32*, 11469−11475.

Nozaki, Y., & Tanford, C. (1971) *J. Biol. Chem. 246*, 2211−2217.

Perutz, M. F. (1970) *Nature 228*, 726−734.

Reynolds, J A., Gilbert, D. B., & Tanford, C. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 2925−2927.

Rezai, A. R., & Esmon, C. T. (1995) *J. Biol. Chem. 270*, 16176−16181.

Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151−176.

Roussel, A., & Cambillau, C. (1989) in *Silicon Graphics Geometry Partner Directory (Fall, 1989)* (Silicon Graphics, Ed.) pp 77−78, Silicon Graphics, Mountain View, CA.

Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.

Segal, I. H. (1975) *Enzyme Kinetics*, pp 77−99, John Wiley and Sons Inc., New York.

Skrzypczak-Jankun, E., Carperos, V. E., Ravichandran, K. G., Tulinsky, A., Westbrook, M., & Maraganore, J. M. (1991) *J. Mol. Biol. 221*, 1379−1393.

Stubbs, M. T., & Bode, W. (1993) *Thromb. Res. 69*, 1−58.

Weber, P. C., Lee, S.-L., Lewandowski, F. A., Schadt, M. C., Chang, C.-H., & Kettner, C. A. (1995) *Biochemistry 34*, 3750−3756.

BI952164B