Kinetic and Crystallographic Studies of Thrombin with Ac-(D)Phe-Pro-boroArg-OH and Its Lysine, Amidine, Homolysine, and Ornithine Analogs

Patricia C. Weber,*,‡ Sheng-Lian Lee,§ Francis A. Lewandowski, Margaret C. Schadt, Chong-Hwan Chang, and Charles A. Kettner

Chemical and Physical Sciences Department, The Du Pont Merck Pharmaceutical Company, P.O. Box 80228, Wilmington, Delaware 19880-0228

Received October 20, 1994; Revised Manuscript Received January 4, 1995⊗

ABSTRACT: The X-ray crystallographic structure of Ac-(D)Phe-Pro-boroArg-OH [DuP714, K_i = 0.04 nM; Kettner, C., Mersinger, L., & Knabb, R. (1990) *J. Biol. Chem. 265,* 18289] complexed with human α-thrombin shows the boron atom covalently bonded to the side-chain oxygen of the active site serine, Ser195. The boron adopts a nearly tetrahedral geometry, and the boronic acid forms a set of interactions with the protein that mimic the tetrahedral transition state of serine proteases. Contributions of the arginine side chain to inhibitor affinity were examined by synthesis of the ornithine, lysine, homolysine, and amidine analogs of DuP714. The basic groups interact with backbone carbonyl groups, water molecules, and an aspartic acid side chain (Asp189) located in the thrombin S1 specificity pocket. The variation in inhibition constant by 3 orders of magnitude appears to reflect differences in the energetics of interactions made with thrombin and differences in ligand flexibility in solution. Kinetic and crystallographic data are reported for the following thrombin inhibitors: DuP714 (space group *C*2, a = 70.8 Å, b = 72.3 Å, c = 72.6 Å, β = 100.6°, crystallographic *R*-factor = 0.204 to 1.95 Å resolution); Ac-(D)Phe-Pro-boroLys-OH (K_i = 0.24 nM, *C*2, a = 70.3 Å, b = 71.9 Å, c = 71.9 Å, β = 100.9°, *R*-factor = 0.201 to 2.35 Å resolution); Ac-(D)Phe-Pro-boro-homoLys-OH (K_i = 8.1 nM, *C*2, a = 70.3 Å, b = 71.9 Å, c = 71.9 Å, β = 100.9°, *R*-factor = 0.212 to 2.4 Å resolution); Ac-(D)Phe-Pro-boroOrn-OH (K_i = 79 nM, *C*2, a = 70.4 Å, b = 72.2 Å, c = 72.2 Å, β = 100.1°, *R*-factor = 0.195 to 2.25 Å resolution); and Ac-(D)Phe-Pro-boro-*n*-butylamidinoGly-OH (K_i = 0.29 nM, *C*2, a = 70.8 Å, b = 72.4 Å, c = 72.2 Å, β = 100.3°, *R*-factor = 0.197 to 2.25 Å resolution).

Many factors, including the formation of favorable van der Waals, hydrogen-bonding, and electrostatic interactions, the entropically disfavored immobilization of ligand and protein on complex formation, and the role of solvent molecules, determine the overall interaction energy between a ligand and its macromolecular target. The extent to which these factors can be estimated by summing the contributions of individual atoms in the ligand is important for understanding protein:ligand interactions in general (Jencks, 1981; Kati et al., 1992) and forms the basis for many computational methods for the analysis of ligand:macromolecular interactions (e.g., Bash et al., 1987; Wilson et al., 1991; Kollman, 1994). Given the complex interplay of affinity-influencing factors such as inhibitor solvation and desolvation (Warshel et al., 1989; Schellman, 1990), the development of predictive computational methods critically depends on detailed binding and structural analyses of ligands or inhibitors having relatively minor chemical changes. Reliable quantitative expectations for the effects that chemical modifications may have on inhibitor affinity can then guide synthesis in a drug discovery effort. Moreover, differences between expected and observed changes in inhibitor affinity for analogs within a chemical series provide feedback regarding the possibility of an alteration in the ligand binding mode (Andrews et al., 1984), a situation typically difficult to detect in the absence of structural analysis.

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 (Stubbs & Bode, 1993). Compounds that inhibit thrombin, such as the 65-residue protein hirudin isolated from the medicinal leech, are effective inhibitors of blood clotting. Clinical advantages exist for the use of low molecular weight thrombin inhibitors in place of the high molecular weight complex of antithrombin III and heparin because such molecules are effective in both blockage of arterial thrombosis (Hanson & Harker, 1988) and venous blood clot accretion (Jackson et al., 1993; Tapparelli et al., 1993). For these reasons, intense research is aimed at the discovery of safe, orally active, synthetic thrombin inhibitors (Fenton et al., 1991; Doorey et al., 1992).

One of the most potent synthetic thrombin inhibitors, H-(D)Phe-Pro-boroArg-OH, binds thrombin with a picomolar inhibition constant (K_i < 4 pM; Kettner et al., 1990). Similar to the interactions of other boronic acid inhibitors with serine proteases (Matthews et al., 1975; Tulinsky & Blevins, 1987; Bone et al., 1987, 1989; Bachovchin et al., 1988; Tsilikounas et al., 1993), part of the high affinity for thrombin is expected to result from the formation of a tetrahedral adduct between the active site serine, Ser195, and the inhibitor boron. Indeed, the crystal structure of the DuP714:thrombin complex, the first boronic acid inhibitor of thrombin studied by X-ray crystallography, shows covalent bond formation between inhibitor and enzyme and provides further evidence that boronic acids bind to serine proteases in a manner that

* Author to whom correspondence should be addressed.
‡ Current address: Schering-Plough Research Institute, 2015 Galloping Hill Road, K-15-3-3855, Kenilworth, NJ 07033-0539.
§ Du Pont Agricultural Products Department, Stine-Haskell Research Center, P.O. Box 30, S300/454, Newark, DE 19714.
⊗ Abstract published in *Advance ACS Abstracts,* March 1, 1995.

0006-2960/95/0434-3750$09.00/0

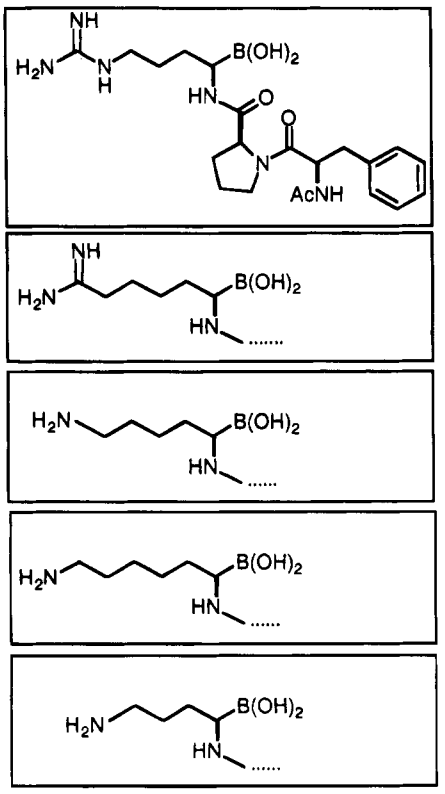


FIGURE 1: Chemical structures of DuP714 (top) and its amidine, lysine, homolysine, and ornithine analogs (second from top to bottom, respectively).

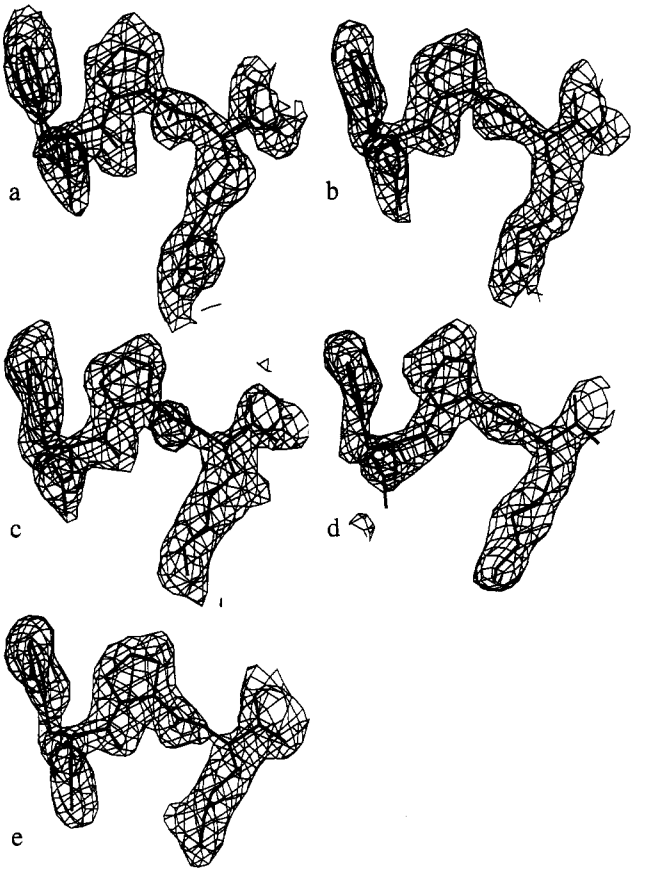


FIGURE 2: ($2F_o - F_c$) electron density maps for DuP714 (a) and its amidine (b), lysine (c), homolysine (d), and ornithine (e) analogs calculated using phases derived from the final model. Contours represent electron density within 2 Å of inhibitor atoms. Density near the boronic acid groups appears jagged because of nearby solvent molecules and covalent bonding to the active site serine (not shown).

mimics the tetrahedral complex formed during normal substrate hydrolysis. An additional contributor to high affinity for this sequence is the packing between the Phe and proline rings that NMR data indicate form in solution prior to complexation with thrombin (Lim et al., 1993). Similar preorganization of these residues may account for the relatively high thrombin affinities found for the aldehyde-containing (Bajusz et al., 1978) and chloromethyl ketone-containing (Kettner & Shaw, 1979) analogs of the (D)Phe-Pro-Arg peptide.

To examine the contributions of the arginyl side chain to the high affinity of DuP714, a number of analogs having basic side chains were synthesized and analyzed kinetically and by X-ray crystallography. Here, we also report the thrombin inhibition constants and crystallographic complex structures of the lysine, homolysine, and ornithine analogs of DuP714 and an amidine analog where the Nϵ nitrogen of arginine was changed to a carbon (Figure 1). Comparison of the X-ray crystallographic structures of each of these inhibitors bound to thrombin revealed an overall similar mode of binding, especially for the Ac-(D)Phe-Pro-boronic acid portion. As discussed here, despite their similarities in chemical structure and in many of their interactions with thrombin, the variation in inhibitor affinity is surprisingly wide and likely reflects energetic differences in the interactions made by the basic groups of these inhibitors, as well as differences in inhibitor flexibility.

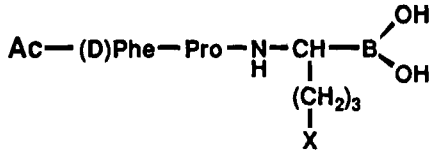
EXPERIMENTAL PROCEDURES

Chemical Synthesis. Peptide analogs of Ac-(D)Phe-Pro-boroArg-OH, where the side chain of the boronic acid was varied, were prepared. The preparation of Ac-(D)Phe-Pro-NH-CH[$(CH_2)_3Br$]-BO_2-$C_{10}H_{16}$ and its conversion to Ac-(D)Phe-Pro-NH-CH[$(CH_2)_3NH_2$]-BO_2-$C_{10}H_{16}$ have been described previously (Kettner et al., 1990). Note that $C_{10}H_{16}$ is the abbreviation for the pinanediol boronic acid ester. Ac-(D)Phe-Pro-NH-CH[$(CH_2)_4Br$]-BO_2-$C_{10}H_{16}$ and Ac-(D)Phe-Pro-NH-CH[$(CH_2)_4NH_2$]-BO_2-$C_{10}H_{16}$ were prepared by an analogous series of reactions from a precursor containing an additional methylene. Ac-(D)Phe-Pro-NH-CH[$(CH_2)_4Br$]-BO_2-$C_{10}H_{16}$ was converted to the corresponding nitrile by treatment with tetra-*n*-butylammonium cyanide, which in turn was reduced by hydrogenation at atmospheric pressure in the presence of 10% Pd/C to yield Ac-(D)Phe-Pro-NH-CH-[$(CH_2)_5NH_2$]-BO_2-$C_{10}H_{16}$. The corresponding amidine was also prepared from the nitrile using the Pinner synthesis, followed by treatment with ammonia (Neilson, 1975). Studies with boroArg and boroLys were conducted with the free boronic acid, while the pinanediol ester was used for the remaining studies. Almost identical behavior is obtained with the ester that is in rapid equilibrium with the free boronic acid. For the two compounds studied as the free boronic acid, the pinanediol group was removed by transesterification with phenylboronic acid (Kettner et al., 1990).

A detailed description of the syntheses along with analytical data is given in the supplementary material.

Enzyme Assays. Inhibitor binding constants were determined at 25 °C in 0.10 mM sodium phosphate buffer (pH 7.5), containing 0.20 M sodium chloride and 0.5% poly(ethylene glycol) MW 6000. Assays were conducted with

Table 1: Summary of Binding and Crystallographic Data for Thrombin:Inhibitor Complexes



	X				
	$NHC(NH)NH_2$	$CH_2C(NH)NH_2$	NH_2	CH_2NH_2	$CH_2CH_2NH_2$
K_i (nM)[a]	0.04[b]	0.29 ± 0.13	79 ± 2	0.24 ± 0.09	8.1 ± 3.6
molecular weight[c]	458.8	457.8	416.8	430.8	444.8
unit cell parameters					
a (Å)	70.8	70.8	70.4	70.3	70.3
b (Å)	72.3	72.4	72.2	71.9	71.9
c (Å)	72.6	72.2	72.2	71.9	71.9
β (deg)	100.6	100.3	100.1	100.9	100.9
R_{sym} (on *I*)	0.077	0.064	0.071	0.064	0.079
resolution range (Å)	6−1.95	6−2.25	6−2.25	6−2.35	5.5−2.4
no. of reflections[d]	20071 (3σ)	12435 (2σ)	11468 (2σ)	12128 (1σ)	8634 (1σ)
crystallographic *R*-factor	0.204	0.197	0.195	0.201	0.212
no. of atoms[e]					
solvent	115	115	121	111	100
ligand	33	33	30	31	32
rms deviation in					
bond distance (Å)	0.014	0.011	0.010	0.011	0.010
angle distance (Å)	0.026	0.022	0.021	0.021	0.019
planar group distance (Å)	0.025	0.019	0.018	0.018	0.017
chiral volume (Å^3)	0.208	0.182	0.173	0.186	0.175
rms difference in[f]					
α-carbons		0.16	0.21	0.22	0.32
all atoms		0.36	0.38	0.38	0.59

[a] K_i values were measured at pH 7.5 and 25 °C as described in the Experimental Procedures section. The homolysine ($CH_2CH_2NH_2$) analog was tested as the free boronic acid, while the remaining compounds were introduced as their corresponding pinanediol esters. K_i values are reported as the averages of at least duplicate measurements with values of standard deviation. [b] Kettner et al., 1990. [c] Molecular weight of the boronic acid form. [d] Reflections having a σ value less than that indicated in the parentheses were eliminated from the crystallographic refinement. [e] Each model contains 2288 protein atoms. [f] Structures were superimposed on the DuP714:thrombin complex using 249 α-carbons of residues 1B−14J of the thrombin A-chain and 16−146 and 150−245 of the B-chain. Deviations are also reported for all 2018 atoms of these residues.

human α-thrombin (specific activity 2340 NIH units/mg) at a concentration of 0.2 nM and the chromagenic substrate H-(D)Phe-Pip-Arg-pNA (S2238; Kettner et al., 1990). K_m values were measured over a substrate range of 20−200 μM. K_i values, determined by allowing the enzyme to react with substrate and inhibitor for 30 min, are given in Table 1. Velocities measured in the time frame of 25−30 min readily fit the equation for competitive inhibition.

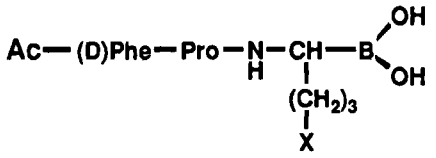
Crystallographic Studies. Human α-thrombin (3.5 mg/mL) was purchased from Enzyme Research Laboratories (South Bend, IN) and used without further purification. Crystals of human α-thrombin having the active site free of inhibitor or substrate were grown by vapor diffusion techniques using conditions described by Skrzypczak-Jankun et al. in 1991. For X-ray diffraction studies of thrombin: inhibitor complexes, 2.0 mg of inhibitor was added to 250 μL of crystal stabilization buffer (0.058 M sodium phosphate (pH 7.2), 33% poly(ethylene glycol) MW 8000, and 0.05 mM NaN_3) either as a solid or as a concentrated DMSO suspension. To allow complete complex formation in the crystalline state, thrombin crystals were soaked in the inhibitor-containing solution for at least 24 h prior to mounting crystals in capillaries for X-ray data collection.

X-ray diffraction data for the thrombin:inhibitor complexes were collected on either a Siemens imaging proportional counter mounted on an Elliott GX-21 rotating anode generator operating at 40 mA and 12 kV or an R-axis Image Plate system mounted on a Rigaku RU-200 rotating anode generator operating at 100 mA and 50 kV. Proportional counter data were reduced to integrated intensities using XENGEN data reduction software (Howard et al., 1987), and image plate data were similarly reduced (Higashi, 1990; Sato et al., 1992). 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 1 lists the crystallographic data parameters for thrombin:inhibitor complex crystals.

Coordinates of the ternary complex of human α-thrombin, the C-terminal hirudin peptide (Skrzypczak-Jankun et al., 1991), and (D)Phe-Pro-Arg-chloromethyl ketone (PPACK; Bode et al., 1992) were used as the starting model for crystallographic refinement by restrained least-squares methods (Hendrickson & Konnert, 1980; Finzel, 1987). For the crystal structure of thrombin complexed with DuP714, positions of solvent molecules were reassigned, and some minor adjustments in the protein and peptide conformations were made, as indicated by the presence of negative density in $(F_o - F_c)\alpha_{calc}$ electron density maps visually inspected at a computer graphics terminal (Jones, 1982). Residues corresponding to the thrombin A-chain from 1B to 14K (thrombin residues are numbered according to homology with chymotrypsinogen; Bode et al., 1989), the B-chain from 1 to 146 and from 150 to 247, and the hirudin peptide from 54 to 60 were clearly defined in the electron density maps.

For thrombin complexed with the DuP714 analogs, refinement at 2.5 Å resolution was initiated using the thrombin: hirudin peptide:solvent model obtained in the refinement of

Table 2: Interactions between Thrombin and Inhibitors



hydrogen-bonding interactions	X					
	$NHC(NH)NH_2$	$CH_2C(NH)NH_2$	NH_2	CH_2NH_2	$CH_2CH_2NH_2$	*a*
Ser195 N–O1	3.2	3.1	2.9	3.2	2.8	2.9
Gly193 N–O1	3.0	3.2	3.3	2.9	3.0	2.6
His57 Nε2–O2	2.6	2.5	2.8	2.7	2.6	2.7
Wat1–O2	2.4	2.3	2.4	2.7	2.7	3.3
Wat1–K60F Nζ	2.9	3.0	3.3	2.5	3.4	
Wat2–O1	2.7	2.4		3.0		3.3
Wat3–O2	2.8	2.7		2.9	2.6	
Wat3–Pro O	3.2	2.9		3.5	2.4	
Gly214 O–boroX N	3.2	3.3	3.2	3.1	2.9	3.0
Gly216 N–(D)Phe O	3.2	3.1	3.0	3.3	2.8	3.0
Gly216 O–(D)Phe N	2.9	2.9	2.7	3.2	2.8	2.9
Asp102 Oδ1–His57 Nδ1	2.8	2.8	2.8	2.8	2.6	2.8
His57 Nε2–Ser195 Oγ	2.7	2.7	3.0	2.7	3.1	3.0
Trp215 Cδ2–(D)Phe Cδ1	3.6	3.4	3.7	3.5	3.3	
Trp215 Cε2–(D)Phe Cε1	4.6	4.2	4.7	4.2	3.9	

angle	X (deg)					
	$NHC(NH)NH_2$	$CH_2C(NH)NH_2$	NH_2	CH_2NH_2	$CH_2CH_2NH_2$	*a*
Ser195 Cβ–Ser195 Oγ–B	134	135	129	130	133	135
Ser195 Oγ–B–O1	104	102	110	110	110	99
Ser195 Oγ–B–O2	101	97	98	101	106	93
Ser195 Oγ–B–boroX Cα	106	124	116	112	118	104
O1–B–O2	104	110	107	107	106	115
O1–B–boroX Cα	119	108	111	108	108	117
O2–B–boroX Cα	120	115	114	118	117	121
B–boroX Cα–boroX N	110	111	110	111	116	111

[a] Bone et al., 1987. Convention for labeling of water oxygen atoms, O1 and O2, was taken from this reference. See also Figure 3.

the DuP714:thrombin complex. Crystallographic refinement was gradually extended to near the limit of diffraction, where the average reflection intensity was nearly twice the magnitude of the background. Models of the inhibitors were built using standard peptide geometry and minimized by steepest descent methods (InsightII, Biosym Technologies, San Diego, CA). In these models, groups bonded to the boron were arranged in a tetrahedral geometry. Inhibitor geometric restraints for refinement were taken from the standard model and incorporated into the refinement program PROFFT (Finzel, 1987) using a program kindly provided by Dr. S. T. Freer. Inhibitor atoms were fitted into the positive electron density of $(F_o - F_c)\alpha_{calc}$ maps calculated using the refined coordinates of the inhibitor-free model. As input to the crystallographic refinement procedure, the boron-to-oxygen distance was restrained to be 1.55 Å. In general, inhibitors were clearly defined in the original omit maps and $(2F_o - F_c)\alpha_{calc}$ electron density maps (Figure 2). Solvent positions and protein and peptide electron density were checked prior to the completion of crystallographic refinement. Final refinement statistics are listed in Table 1.

RESULTS

Inhibition Constants. Although DuP714 and its analogs (Figure 1) share the Ac-(D)Phe-Pro-boronic acid sequence attached to a basic side chain, the thrombin inhibition constants varied by 3 orders of magnitude from 79 to 0.040 nM (Table 1). Reaction velocities, measured 25–30 min after thrombin was mixed with substrate and inhibitor, readily fit the equation for competitive inhibition. While DuP714 is a slow-binding inhibitor of thrombin (Kettner et al., 1990), the remaining inhibitors whose affinities for thrombin are lower did not appear to exhibit significant slow-binding properties. No greater than a 3-fold difference was observed in comparing the initial and steady state velocities.

Boronic Acid:Thrombin:Solvent Interactions. The boronic acid group of these inhibitors is extensively bonded to thrombin and associated water molecules (Table 2). Indicative of bond formation, the electron density between the boron atom of the inhibitor and the side-chain oxygen of Ser195 is continuous. While the boron atom does become tetracoordinate upon complex formation, as expected for this mimic of the transition state, its geometry is more trigonal than that of the tetrahedral model used to generate the crystallographic restraints. Lack of perfect tetrahedral geometry was also apparent by inspection of the electron density maps at various stages of the crystallographic refinement process, including the original maps calculated before the inhibitor was included in the model and the final electron density maps. Slightly distorted tetrahedral geometries have also been observed for the boronic acid groups of peptide-like inhibitors of α-lytic protease (Bone et al., 1987, 1989), although in the complexes of subtilisin with benzeneboronic acid (Matthews et al., 1975) and of α-chymotrypsin with 2-phenylethaneboronic acid (Tulinsky &

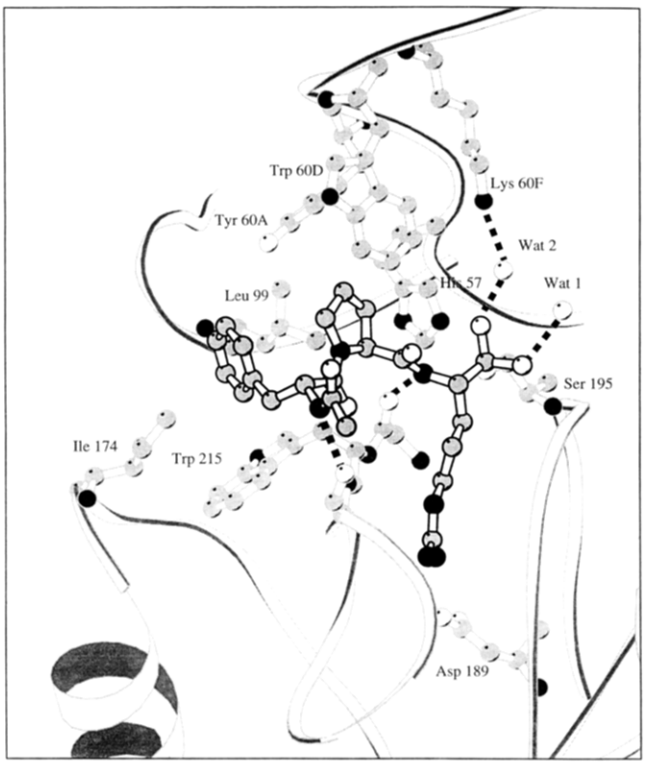


FIGURE 3: Schematic representation of DuP714 bound at the active site of thrombin. A narrow ribbon (Kraulis, 1991) outlines the thrombin fold. Shaded circles indicate carbon atoms, white circles oxygen atoms, and solid circles nitrogen atoms. Solid bonds interconnect DuP714 atoms, and dashed lines indicate hydrogen bonds. Inhibitor:thrombin interactions are described further in the text.

Blevins, 1987) the boron adopts a tetrahedral geometry. This suggests that the geometry about the boron atom is influenced by the other interactions made between inhibitor and protease. For smaller inhibitors, the boron can assume a tetrahedral geometry, whereas for the larger, peptide-like inhibitors, whose interactions with thrombin are more extensive, the lack of tetrahedral geometry reflects a compromise between the optimization of boron:active site interactions and other inhibitor:protein interactions.

The boronic acid hydroxyl groups are extensively hydrogen bonded to both the protein and the solvent (Figure 3, Table 1). One interacts with backbone amide hydrogen bond donors of the oxyanion hole (Gly193 and Ser195; Kraut, 1977) and with a water molecule hydrogen bonded to the backbone carbonyl of Leu41. The other is hydrogen bonded to the His57 side chain and a water molecule. Taken together, the short boron-to-Ser195 Oγ distance, the nearly tetrahedral geometry about the boron, and the stabilization of boronic acid binding by hydrogen bonding to the oxyanion hole suggest that the carboxy terminal boronic acid group acts as a mimic of the transition state of serine proteases (Kraut, 1977).

General Features of Thrombin:Inhibitor Interactions. The Ac-(D)Phe-Pro portions of these inhibitors bind with the tripeptide backbone in an extended conformation and form three hydrogen bonds (Table 2) with thrombin main-chain β-sheet atoms located at the base of the substrate binding site (Figure 3). The hydrophobic proline ring adopts a nearly planar conformation and packs in a narrow pocket lined on either side by the side chains of Tyr60A and Trp60D. The aromatic ring of the (D)Phe residue is packed into a relatively hydrophobic pocket lined by the side chains of Leu99 and Ile174 and the main-chain carbonyl oxygen atoms of Asn98 and Leu99. The (D)Phe aromatic ring forms a favorable edge-to-face stacking interaction (Cox et al., 1958; Janda et al., 1975) with the side chain of Trp215 located at the base of the pocket. Similar to the interaction observed by solution NMR (Lim et al., 1993), the (D)Phe aromatic ring is also within van der Waals contact of the proline ring, suggesting that the high affinity of this tripeptide sequence may arise, in part, because of its propensity to adopt similar packing arrangements in solution and when bound to thrombin.

Overall, the packing interactions and hydrogen bonds made by DuP714 with thrombin are similar to those observed in the complex between PPACK and thrombin (Bode et al., 1992). However, the electronic differences in the chloromethyl ketone (CMK) and boronic acid groups result in subtle differences in their interactions with thrombin. Unlike the boronic acid, which when bound to the Ser195 side chain acts as a transition state mimic, the CMK group of PPACK is covalently linked to the protein by the formation of a hemiketal adduct, which involves bonding to both the Ser195 and His57 side chains (Bode et al., 1992). As a result of its covalent interaction with only Ser195 of the catalytic triad, the entire DuP714 molecule is shifted slightly toward the active site serine and the oxyanion hole relative to PPACK. The shift is sufficiently small, so that the packing interactions between thrombin and the inhibitor proline and Phe aromatic rings and the charge−charge interaction between inhibitor guanidino and the Asp189 of thrombin (described in the following) are present in both complexes.

Basic Side Chain:Thrombin:Solvent Interactions. The basic P1 side chain of each inhibitor binds in a relatively extended conformation in the thrombin S1 specificity pocket (Figure 3). However, differences in the chemical nature of the basic group and the composition and length of the spacer between it and boronic acid result in variations in binding interactions and the role of solvent in stabilization of inhibitor:thrombin interactions.

The guanidino side chain of DuP714 forms a bidentate hydrogen-bonding interaction with the side chain of Asp189 located at the bottom of the S1 specificity pocket. As shown in Figure 4a, the terminal nitrogens, NH1 and NH2, are roughly equidistant from the side-chain carboxylic oxygens of Asp189. Each NH nitrogen is additionally hydrogen bonded to the protein, one through a water-mediated hydrogen bond to the backbone carbonyl of Phe227 and the other to the backbone carbonyl of Gly219. The guanidino Nϵ nitrogen is hydrogen bonded through a water molecule to the backbone carbonyl of Gly219. The terminal nitrogens of the amidine analog are similarly hydrogen bonded to both the Asp189 side chain and the carbonyl oxygens of Phe227 and Gly219. However, substitution of a carbon for the Nϵ nitrogen of the guanidino group results in the loss of the water-mediated hydrogen bond between the inhibitor side chain and Gly219 backbone carbonyl.

The homolysine side chain is also directly bonded to the Asp189 with the terminal nitrogen about 3 Å from each carboxylate oxygen. The terminal nitrogen also forms longer hydrogen bonds (3.6 and 4.1 Å, respectively) to the backbone carbonyl oxygens of Ala190 and Gly219. A weak, water-mediated hydrogen bond is made between a solvent molecule (Wat4, Figure 4d) that is 3.8 Å from the homolysine terminal

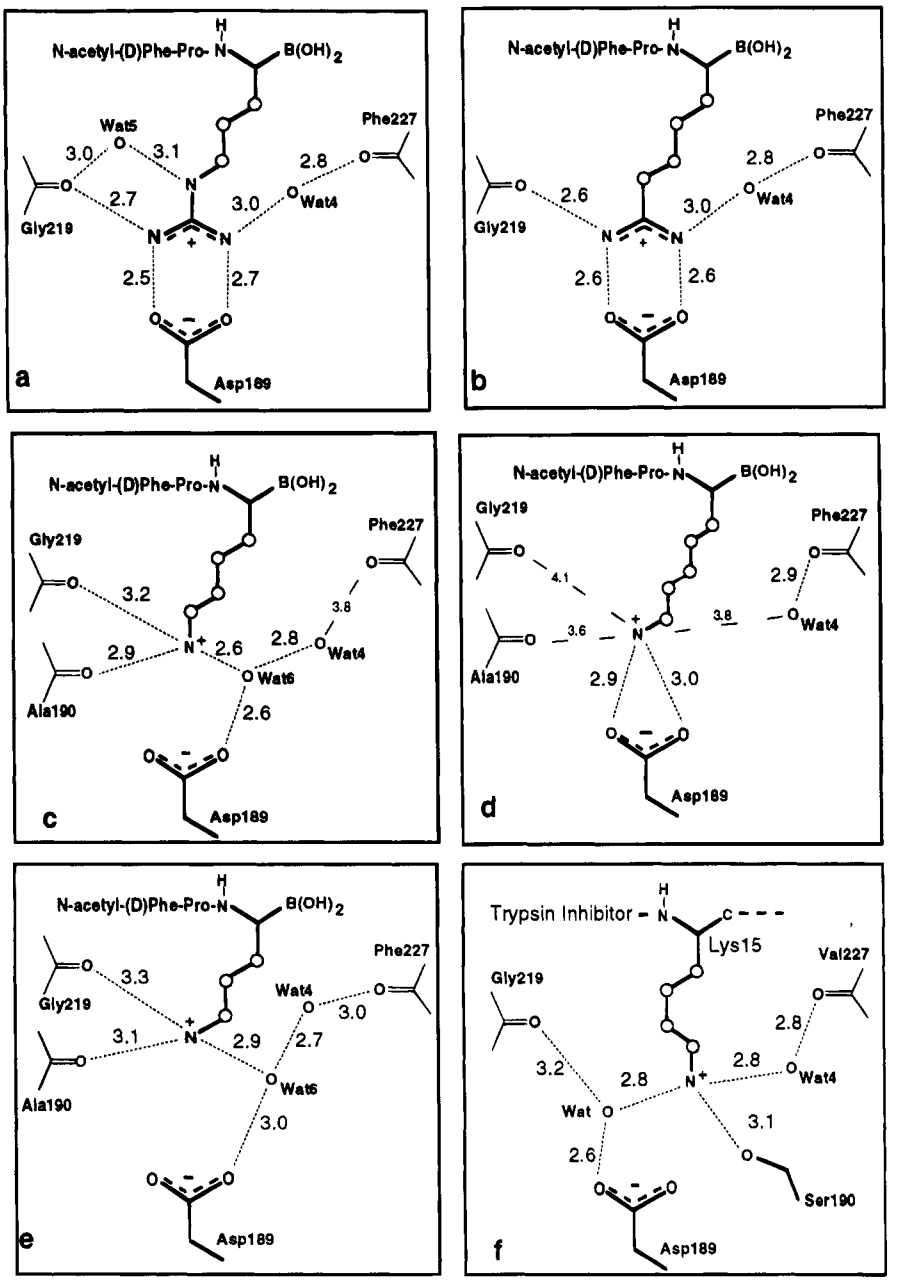


FIGURE 4: Schematic illustration of interactions formed by DuP714 (a) and its amidine (b), lysine (c), homolysine (d), and ornithine (e) analogs in the S1 specificity pocket of thrombin. Dashed lines indicate hydrogen bonds and charge−charge interactions. Distances (in angstroms) are also shown. For the lysine analog (c), the distance between the Nε nitrogen and the Asp189 side-chain oxygen not involved in the water-mediated hydrogen bond is 3.5 Å, and the distance to the other side-chain oxygen is 3.9 Å. For the ornithine analog (e), these distances are 4.3 and 4.5 Å, respectively. The terminal nitrogen atom of the homolysine side chain (d) is within 4.1 Å of the backbone carbonyl of Gly219, 3.6 Å of the backbone carbonyl of Ala190, and 3.8 Å of an immobilized water molecule (Wat4). Part f shows the interactions made by the lysine side chain of bovine pancreatic trypsin inhibitor bound in the S1 specificity pocket of trypsin (Ruhlmann et al., 1973). Comparison of parts c and f shows that the water-mediated interactions between inhibitor lysine side chains involve different side-chain oxygen atoms in trypsin and thrombin.

nitrogen and 2.9 Å from the backbone carbonyl oxygen of Phe227.

Unlike the guanidino, amidine, and homolysine analogs, neither the lysine nor ornithine side chains interact directly with Asp189. Instead, these basic side chains are hydrogen bonded to both solvent molecules and backbone carbonyl groups (Figure 4c,e). The interaction between terminal nitrogen and the Asp189 carboxylate is mediated by a water molecule, which along with an additional solvent molecule participates in a hydrogen bond network with the backbone carbonyl oxygen of Phe227. Similar to the homolysine case, the lysine and ornithine terminal nitrogens are hydrogen bonded to the backbone carbonyl of Ala190. A water-mediated interaction between the inhibitor lysine side chain and the protease aspartate side chain has also been observed in the complexes of trypsin with bovine pancreatic trypsin inhibitor (Ruhlmann et al., 1973) and (carbobenzyloxy)lysl chloromethyl ketone (Marquart et al., 1983). The substitution of Ala190 in thrombin for a serine in trypsin results in a subtle difference in the interaction of inhibitor lysine side chains within the S1 specificity pockets of thrombin and trypsin. The lysine side chain of the trypsin inhibitors is hydrogen bonded to the side-chain oxygen of Ser190, not the backbone carbonyl of the residue at this position, as in

the case with thrombin. This alteration in hydrogen bonding from the backbone to the side chain of residue 190 situates the lysine side chain of the trypsin inhibitors more deeply in the protein interior, and in addition to other hydrogen-bonding changes diagrammed in Figure 4c,f, the water-mediated hydrogen bond between the inhibitor lysine and protein aspartate involves opposite aspartate oxygens in thrombin and trypsin.

DISCUSSION

An objective of this study was to determine how relatively conservative changes in the basic side chain of DuP714 would influence its thrombin binding properties. Taken together, the binding and crystallographic data show that despite changes that alter inhibitor molecular weights by only 42, whose variation in chemical structure is confined to a single part of the molecule and whose variation in chemical structure retains a positively charged moiety that in the thrombin:inhibitor complex is similarly oriented in the S1 specificity pocket, the affinity for thrombin changed by a factor of nearly 2000. While the enumeration of all factors and their relative energetic contributions to account for this surprising variation in affinity is difficult, several that may contribute to the measured differences can be identified.

Increases in inhibitor flexibility can decrease affinity. Here, several examples, where minor changes in chemical structure resulted in side chains expected to be more flexible in solution, are presented. For example, side-chain flexibility was increased by substitution of the guanidino Nϵ nitrogen of DuP714 for a carbon in the amidine analog and by the addition of a methylene group into the lysine side chain to form the homolysine analog. In both examples, the similar side-chain *B*-values observed in the crystal structures of the inhibitor:thrombin complexes suggest that entropically disfavored side-chain immobilization contributes to the decreased affinity of the more flexible molecules.

Loss of hydrogen-bonding capability can decrease inhibitor affinity. The greater than 6-fold difference in affinity between the guanidino and amidine analogs may be due, in part, to a decrease in the number of hydrogen bonds formed between inhibitor and protein.

Bidentate interactions between oppositely charged groups are more stable than monodentate interactions. Consistent with a number of ligand binding studies that show that the higher affinity of compounds capable of forming bidentate interactions, whose enhanced stabilization arises from their ability to form alternate, nearly isoenergetic hydrogen-bonding patterns (Sapse & Russell, 1984), the stronger thrombin complex formed by the amidine analog compared to the homolysine may reflect the ability of the former to make an energetically more favorable bidentate charge–charge interaction with the aspartate side chain at the bottom of the thrombin S1 specificity pocket.

Longer hydrogen bonds can decrease inhibitor affinity. The same groups participate in hydrogen bonding to both the lysine and ornithine analogs (Figure 4c,e). The increased affinity of the lysine analog relative to the ornithine can be due to the fact that all of the hydrogen bonds are shorter in the lysine analog complex.

Increased basicity may promote binding. The guanidino group has a higher pK_a (13.6 vs 12.4 for the amidine analog; Patel 1975, 1991), so that, at the same pH, a greater fraction of the guanidino-containing inhibitors is fully charged than for the corresponding amidine analog. This difference in charge, although slight at the assay pH, may also contribute to the higher affinity of the guanidino-containing compound.

Bound solvent molecules can provide excellent interaction sites for inhibitors. Given all of the preceding pairwise comparisons, the lysine analog has an unusually high affinity for thrombin. This analog interacts with the thrombin aspartate only through a water-mediated hydrogen bond, while the homolysine analog, which binds with a 33-fold lower affinity, interacts directly. One might have expected that the homolysine analog, which in addition to forming the direct electrostatic interaction also displaces an additional bound water molecule (Wat6, Figure 4), might form the tighter complex with thrombin. However, as mentioned earlier, relative to lysine, the binding of the homolysine side chain involves the immobilization of an additional methylene group. It appears that the entropically favorable displacement of water from the S1 specificity pocket by the homolysine may be offset by unfavorable side-chain immobilization. In protein engineering studies of rat trypsin, an unexpected high affinity for a water-mediated lysine-to-aspartate interaction has also been observed (Perona et al., 1993).

In conclusion, for DuP714 and the analogs studied here, seemingly minor modifications in inhibitor structure produced large changes in binding affinity, despite the preservation of many interactions with the target enzyme, as shown by crystallographic structures of the inhibitors complexed with thrombin. Taken together, these studies again demonstrate the delicate balance between many individual enthalpic and entropic factors that contribute to the overall affinity between inhibitor and protein.

SUPPLEMENTARY MATERIAL AVAILABLE

Detailed description of the chemical syntheses along with analytical data for compounds used in this study (5 pages). Ordering information is given on any current masthead page.

REFERENCES

Andrews, P. R., Craik, D. J., & Martin, J. L. (1984) *J. Med. Chem. 27*, 1648.

Bachovchin, W. W., Wyl-Wong, B., Farr-Jones, S., Shenvi, A. B., & Kettner, C. A. (1988) *Biochemistry 27*, 7689.

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

Bash, P. A., Singh, U. C., Brown, F. K., Langridge, R., & Kollman, P. A. (1987) *Science 235*, 574.

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

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

Bone, R., Shenvi, A. B., Kettner, C. A., & Agard, D. A. (1987) *Biochemistry 26*, 7609.

Bone, R., Frank, D., Kettner, C. A., & Agard, D. A. (1989) *Biochemistry 28*, 7600.

Cox, E. G., Cruikshank, D. W. J., & Smith, J. A. (1958) *Proc. R. Soc. London A247*, 1.

Doorey, A. J., Michelson, E. L., & Topol, E. J. (1992) *J. Am. Med. Assoc. 268*, 3108.

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

Finzel, B. C. (1987) *J. Appl. Crystallogr. 20*, 53.

Hanson, S. R., & Harker, L. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 3184.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Biomolecular Structure, Function, Conformation and Evolution* (Srinivasan, R., Ed.) p 43, Oxford University Press, Oxford, U.K.
Higashi, T. (1990) *J. Appl. Crystallogr. 23*, 253.
Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., & Salemme, F. R. (1987) *J. Appl. Crystallogr. 20*, 383.
Jackson, C. V., Wilson, H. C., Growe, V. G., Shuman, R. T., & Gesellchen, P. D. (1993) *J. Cardiovasc. Pharmacol. 21*, 587.
Janda, K. C., Hemminger, J. C., Winn, J. S., Novick, S. E., Harris, S. J., & Klemperer, W. J. (1975) *J. Chem. Phys. 63*, 1419.
Jencks, W. P. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4046.
Jones, T. A. (1982) in *Computational Crystallography* (Sayre, D., Ed.) p 303, Clarendon Press, Oxford, U.K.
Kati, W. M., Acheson, S. A., & Wolfenden, R. (1992) *Biochemistry 31*, 7356.
Kettner, C., & Shaw, E. (1979) *Thromb. Res. 14*, 969.
Kettner, C., Mersinger, L., & Knabb, R. (1990) *J. Biol. Chem. 265*, 18289.
Kollman, P. (1994) *Curr. Opin. Struct. Biol. 4*, 240.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946.
Kraut, J. (1977) *Annu. Rev. Biochem. 46*, 331.
Lim, M. S. L., Johnston, E. R., & Kettner, C. A. (1993) *J. Med. Chem. 36*, 1831.
Marquart, M., Walter, J., Deisenhofer, J., Bode, W., & Huber, R. (1983) *Acta Crystallogr. B39*, 480.
Matthews, D. A., Alden, R. A., Birktoft, J. J., Freer, S. T., & Kraut, J. (1975) *J. Biol. Chem. 250*, 7120.
Neilson, D. G. (1975) in *The Chemistry of Amidines and Imidates*, p 385, John Wiley & Sons, New York.
Patel, S. (1975) in *The Chemistry of Amidines and Imidates*, p 12, John Wiley & Sons, New York.
Patel, S. (1991) in *The Chemistry of Amidines and Imidates*, p 77, John Wiley & Sons, New York.
Perona, J. J., Tsu, C. A., McGrath, M. E., Craik, C. S., & Fletterick, R. J. (1993) *J. Mol. Biol. 230*, 934.
Ruhlmann, A., Kukla, D., Schwager, P., Bartels, K., & Huber, R. (1973) *J. Mol. Biol. 77*, 417.
Sapse, A. M., & Russell, C. S. (1984) *Int. J. Quant. Chem. 26*, 91.
Sato, M., Yamamoto, M., Imada, K., Katsube, Y., Tanaka, N., & Higashi, T. (1992) *J. Appl. Crystallogr. 25*, 348.
Schellman, J. A. (1990) *Biophys. Chem. 37*, 121.
Skrzypczak-Jankun, E., Carperos, V., Ravichandran, K. G., Tulinsky, A., Westbrook, M., & Maraganore, J. M. (1991) *J. Mol. Biol. 221*, 1379.
Stubbs, M. T., & Bode, W. (1993) *Thromb. Res. 69*, 1.
Tapparelli, C., Metternich, R., Ehrhardt, C., & Cook, N. S. (1993) *Trends Pharm. Sci. 14*, 366.
Tsilikounas, E., Kettner, C. A., & Bachovchin, W. W. (1993) *Biochemistry 32*, 12651.
Tulinsky, A., & Blevins, R. A. (1987) *J. Biol. Chem. 262*, 7737.
Warshel, A., Aqvist, J., & Creughton, S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5820.
Wilson, C., Mace, J. E., & Agard, D. A. (1991) *J. Mol. Biol. 132*, 495.

BI942454C