Crystal Structures of Thrombin Complexed to a Novel Series of Synthetic Inhibitors Containing a 5,5-*trans*-Lactone Template

Harren Jhoti,* Anne Cleasby, Scott Reid,‡ Pam J. Thomas, Malcolm Weir, and Alan Wonacott

Biomolecular Structure Unit, Glaxo Wellcome Medicines Research Centre, Gunnels Wood Road, Stevenage, Hertfordshire SG1 2NY, U.K.

Received December 23, 1998; Revised Manuscript Received March 24, 1999

ABSTRACT: The binding modes of four active site-directed, acylating inhibitors of human α-thrombin have been determined using X-ray crystallography. These inhibitors (GR157368, GR166081, GR167088, and GR179849) are representatives of a series utilizing a novel 5,5-*trans*-lactone template to specifically acylate Ser195 of thrombin, resulting in an acyl complex. In each case the crystal structure of the complex reveals a binding mode which is consistent with the formation of a covalent bond between the ring-opened lactone of the inhibitor and residue Ser195. Improvements in potency and selectivity of these inhibitors for thrombin are rationalized on the basis of the observed protein/inhibitor interactions identified in these complexes. Occupation of the thrombin S2 and S3 pockets is shown to be directly correlated with improved binding and a degree of selectivity. The binding mode of GR179849 to thrombin is compared with the thrombin/PPACK complex [Bode, W., Turk, D., and Karshikov, A. (1992) *Protein Sci. 1*, 426−471] as this represents the archetypal binding mode for a thrombin inhibitor. This series of crystal structures is the first to be reported of synthetic, nonpeptidic acylating inhibitors bound to thrombin and provides details of the molecular recognition features that resulted in nanomolar potency.

Thrombin is a serine protease belonging to the trypsin family, and its function is central to the maintenance of hemostasis and control of thrombosis. Its role in the blood coagulation cascade is to generate clots by cleaving fibrinogen, and the resulting fibrin polymerizes to form a matrix for the nascent clot (*1*). Generation of α-thrombin results primarily from cleavage of the zymogen precursor prothrombin by factor Xa which is part of the prothrombinase complex (*2*). Thrombin is also a major activator of platelet aggregation; proteolytic cleavage of the N-terminal extracellular domain of the thrombin receptor by thrombin creates a new N-terminus that serves as a tethered ligand which activates the receptor (*3*). Platelet activation by thrombin leads to shape change, platelet aggregation, and release of secretory granules. Thrombin's enzymatic activity is regulated in vivo by endogenous inhibitors such as α_2-macroglobulin, antithrombin III, heparin cofactor II, and proteinase nexin I. Other potent and specific proteinacous inhibitors of thrombin include hirudin, the anticoagulant found in the saliva of the medicinal leech (*4*), and rhodniin, which is found in the blood-sucking insect *Rhodnius prolixus* (*5*). Clearly, thrombin has a central position in maintaining hemostasis, and as such it has become a key target for treatment of thromboembolic disease (*6*).

Thrombin has been the focus of intensive biochemical and structural studies which have resulted in a detailed description of the structure and function of this serine protease (for review see refs *7* and *8*). The crystal structure of thrombin, like trypsin, reveals two β-barrels with the active site catalytic triad (Ser195, His57, and Asp102)[1] located at the junction of these supersecondary structural elements. In comparison with trypsin, the thrombin active site cleft is deeper and narrower due to several sequence insertions around this region resulting in longer loops that enclose the active site. Two pronounced insertion loops that partly form the thrombin active site and contribute to its restricted substrate specificity are the "60-insertion loop" which includes residues Tyr60a, Trp 60d, and Lys 60f and the "148-insertion loop" which includes Trp148 (*10*). Another striking feature of the thrombin structure is its pronounced electrostatic charge distribution, which is largely due to a positively charged "anion-binding" site or exosite (also known as the fibrinogen recognition exosite) which is remote from the active site and is involved in binding fibrinogen, hirudin, and the thrombin receptor. In addition to this fibrinogen recognition site there is also another positively charged region on the surface of thrombin which has been shown to interact with heparin and, hence, is known as the "heparin-binding site" (*8*).

Thrombin is similar to trypsin in that it exhibits a preference for basic residues at the P1 position of substrates (preferring Arg to Lys); this is reflected in the nature of most of its synthetic substrates and inhibitors (*11*). This preference is due to the presence of the negatively charged residue Asp189 at the base of the S1 pocket.[2] Small hydrophobic residues such as proline are preferred at position P2, the size constraint being due to the relatively small S2 pocket which

‡ Present address: AxyS Pharmaceutical Inc., 180 Kimball Way, South San Francisco, CA 94065.

[1] The amino acid residue numbering follows the chymotrypsinogen system (*9*).

[2] Enzyme pockets (S1, S2, ...) corresponding to substrate/inhibitor residues (P1, P2, ...) follow Schechter and Berger notation (*12*).

10.1021/bi9830359 CCC: $18.00
Published on Web 05/28/1999

is partially defined by the side chains of Tyr60a and Trp60d from the 60-insertion loop. This pocket is also referred to as the P-pocket, "proximal" to the active site serine (*13*), and provides a significant contribution to the narrow substrate specificity of thrombin. A region of thrombin that has been described as the "aryl-binding" pocket is enclosed by the side chains of Trp215, Leu99, Ile174, and Tyr60a. This pocket is also referred to as S3 due to the binding of the P3 fragment of peptide-like inhibitors such as PPACK[3] (which has a D-amino acid at P3) and also as the D-pocket as it is "distal" to the active site serine residue (*13*). Insights into the molecular basis of thrombin's specificity have also come from structural and biochemical data on thrombin in complex with the protein hirudin (*14*), hybrids of active site and exosite hirudin peptides (*15*) and fibrinogen (*16, 17*).

The search for small, orally bioavailable, nonpeptidic, synthetic inhibitors of thrombin has been particularly intense in many pharmaceutical companies. The binding modes as determined by X-ray crystallography of some of these inhibitor series have been described: a 5-amidinoindole lead compound (*18*), a series containing rigid bi- and tricyclic core templates (*19*), compounds containing a pyridinone core (*20*), and a series with 4-aminopyridine as P1 substituents (*21*). These compounds can be described as competitive, reversible "nonelectrophilic" inhibitors that derive their binding energy solely from noncovalent interactions with the protein. Many other studies have described thrombin inhibitors that utilize an electrophilic center as a target for a nucleophilic attack by the enzyme, usually by the active site serine. The mode of action of these "electrophilic" inhibitors results in the formation of a covalent bond between the inhibitor and the serine residue which is usually irreversible (*22*).

As part of an effort to identify novel, small molecule inhibitors of thrombin, we have previously reported the discovery of a class of thrombin inhibitor that was isolated from *Lantana camara* (commonly known as "wild sage") during high-throughput screening (*23*). This novel class of serine protease inhibitor contains a strained 5,5-*trans*-fused lactone which was shown to be responsible for its inhibitory activity. Subsequent mechanistic and biophysical studies established the mode of action of these euphane triterpene molecules to involve acylation of the catalytic Ser195 in the thrombin active site via the strained lactone ring (*24*). However, as these natural product inhibitors lacked significant specificity and were synthetically complex, a chemistry effort was initiated to exploit the novel *trans*-lactone template in the design of simpler synthetic inhibitors for thrombin. A crystal structure of thrombin complexed to one of these compounds (GR133686) provided the basis for molecular modeling studies which allowed potency and selectivity of this series to be optimized (*24*). In this study we present the results of this inhibitor design process by describing the binding modes of some representatives from this series in complex with thrombin as defined by X-ray crystallography. These crystal structures are the first to be determined of thrombin complexed with synthetic, nonpeptidic electrophilic inhibitors, and they identify the key protein/inhibitor interactions that resulted in nanomolar potency for thrombin.

[3] Abbreviations: DMSO, dimethyl sulfoxide; PEG, poly(ethylene glycol); ESI-MS, electrospray ionization mass spectrometry; PPACK, D-Phe-L-Pro-L-Arg chloromethyl ketone.

EXPERIMENTAL PROCEDURES

Materials. The freeze-dried human prothrombin complex intermediate was purchased from Bio Products Laboratory, Dagger Lane, Elstree, U.K. *Oxyuranus scutellatus* snake venom was purchased from Sigma. The heparin−Sepharose CL-6B resin was from Pharmacia, PEG 6000 from Fluka, Tris-HCl from BDH, and sodium chloride from Fisons. All buffers were made up with Milli-Q grade water.

Activation and Purification of α*-Thrombin*. Human α-thrombin was activated and purified following a modified procedure of Ngai and Chang (*25*). One vial of the prothrombin complex intermediate was resuspended at room temperature in a buffer of 20 mM Tris HCl and 0.1 M NaCl at pH 7.5 and diluted to a concentration of 0.2−0.4 mg/mL, estimated using an exctinction coefficient at 280 nm of $\epsilon = 5.2 \times 10^4\ M^{-1}\ cm^{-1}$. In the presence of 10 mM $CaCl_2$, 10 mg of freeze-dried snake venom was dissolved in 1 mL of water and added to the prothrombin under constant stirring. The activation was monitored by the increase in absorbance at 405 nm following the release of nitroaniline from the cleavage of the chromogenic substrate S2238 (*26*). Once the rate of absorption change became constant, the activated solution was placed on ice; activation was usually complete in 30−45 min. The activated α-thrombin was loaded onto a heparin−sepharose column equilibrated with a buffer of 20 mM Tris-HCl and 0.1 M NaCl, pH 7.5, at a flow rate of 5 mL/min. After being loaded, the column was washed until all the unbound inactive material was eluted. The activated α-thrombin was eluted from the column using a linear gradient from 0 to 100% 20 mM Tris HCl and 1 M NaCl, pH 6.0, and pooled fractions were stored at −70 °C.

Crystallographic Studies. Crystals of human α-thrombin complexed to hirugen were prepared by a modification of the method described by Skrzypczak-Jankun et al. (*27*). Typically, crystals grew to $200 \times 150 \times 200\ \mu m^3$, and these were then used in soaking experiments with inhibitors to obtain active site complexes. For example, a crystal would be harvested into a solution containing 30% PEG 4000, 100 mM Hepes, 200 mM NaCl at pH 7.0, and 100−200 μM inhibitor and soaked for 24−48 h at 4 °C prior to data collection. In each case the inhibitors were dissolved in 100% DMSO to produce stock solutions prior to addition to the harvesting solution. The crystals were then capillary mounted and X-ray data collected from one crystal using a FAST area detector system mounted on an Enraf-Nonius FR581 rotating anode. The data sets were processed using MADNES (*28*) and CCP4 programs (*29*). Table 1 contains the crystallographic parameters for the four thrombin/inhibitor complexes. All inhibitors were synthesized as described by Pass et al. (*30*).

Models for the inhibitors were constructed using QUANTA (Molecular Simulations Inc., San Diego, CA) and fitted to difference electron density maps using interactive computer graphics. All refinement procedures were performed using X-PLOR (*31*), and manual refitting and inclusion of water molecules were carried out on a computer graphics terminal using QUANTA. In each case the starting model for crystallographic refinement was derived from the native thrombin/hirugen complex with the water molecules around the active site omitted. This was then used for phasing the initial $(|F_o| - |F_c|)\alpha_{calc}$ maps which were interpreted for

Table 1: Statistics for X-ray Data Collection and Refinement for the Thrombin/Inhibitor Complexes[a]

	complex A GR157368	complex B GR166081	complex C GR167088	complex D GR179849
unit cell				
a (Å)	71.7	71.7	71.8	71.4
b (Å)	72.0	71.9	72.0	72.1
c (Å)	73.0	73.0	73.2	72.8
β (deg)	101.0	100.9	101.0	100.7
R_{merge} (%)	5.8	6.1	4.4	5.9
resolution (Å)	30−2.0	38−2.0	39−2.0	30−2.2
no. of reflections	21972	21478	23518	18523
R-factor (%)	20.4	18.9	20.0	18.7
no. of atoms	2696 (205, 16)	2729 (254, 27)	2736 (233, 28)	2727 (219, 33)
average *B* (Å^3)	28.6 (46.3, 23.1)	32.3 (53.2, 51.4)	32.7 (47.6, 40.9)	33.5 (54.4, 26.1)
occupancy	0.87	0.89	0.75	0.64
rms deviations				
bonds (Å)	0.016	0.016	0.017	0.019
angles (deg)	3.1	3.0	3.2	3.0

[a] The numbers given for R_{merge} and reflections are for all data; no sigma cutoff was applied. Figures in parentheses are for solvent and inhibitor, respectively, and the occupancy is for the inhibitor atoms only.

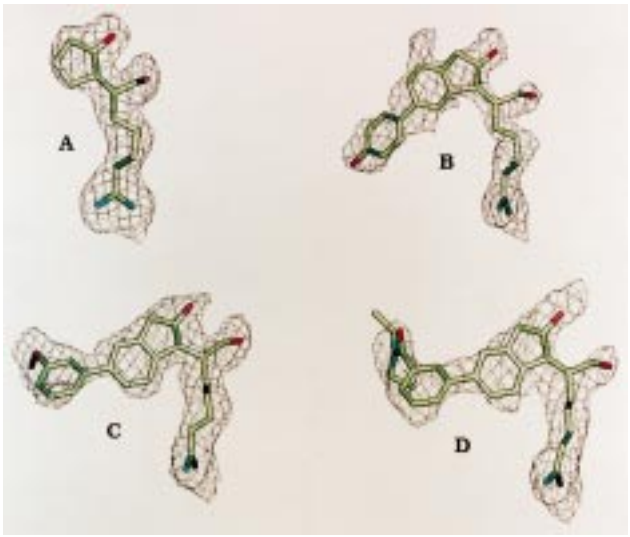


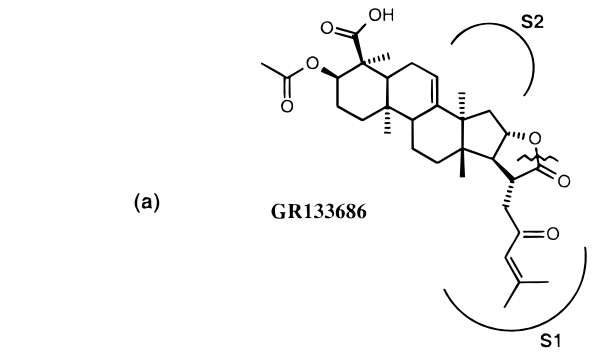
FIGURE 1: Difference electron density for each inhibitor molecule calculated after the refinement shown with the final conformation of the inhibitor. Phases for the $(|F_o| - |F_c|)\alpha_{calc}$ maps were calculated after the inhibitor molecule was omitted from the model, and the maps were contoured at 2.5 σ.

fitting the inhibitor models. After the inhibitor molecule was modeled into the difference electron density, the geometry around the acylated Ser195 residue was restrained to produce a planar ester linkage to the inhibitor. Final refinement and analysis statistics of the four complexes are provided in Table 1.

RESULTS

Structure determination for each complex was trivial as the crystals were isomorphous to the native thrombin/hirugen complex that we had previously solved in our laboratory. In most cases there were strong density features at the thrombin active site in the difference electron density maps; however, this was not the case for GR179849 (complex D) (Figure 1). Initial attempts at obtaining a structure of thrombin complexed to GR179849 resulted in no significant difference electron density in the active site. We presumed this outcome to be a result of hydrolysis of the inhibitor by thrombin and tested this hypothesis by varying the pH during the soaking stage. The harvesting solution was maintained at pH 8.2 for 2 h to optimize binding and then gradually replaced by an identical solution with a pH of 6.0 which was intended to arrest any hydrolysis of the inhibitor and capture the acyl complex. Only after employing this modified soaking protocol were we able to detect significant difference electron density in the thrombin active site indicating that GR179849 was bound. This was confirmed by electrospray ionization mass spectrometric analysis of the crystal after data collection that indicated 64% occupancy of thrombin active sites by GR179849 (data not shown). Despite the significant electron density at the active site for GR166081 (complex B), the interpretation was problematic in that there were features that could not be completely accounted for by the proposed inhibitor conformation (Figure 1). We concluded that, although alternative inhibitor conformations may also be possible, it was most likely that these unexplained density peaks were due to other entities, such as solvent, bound to thrombin. This conclusion was consistent with the electrospray ionization mass spectrometric analysis of the crystal after data collection that indicated only 40% of the thrombin active sites were bound by GR166081 (data not shown). We are, therefore, convinced that the bulk of the difference density is due to the presence of bound GR166081 and that these unexplained density peaks are a consequence of low occupancy of these sites by the inhibitor.

Thrombin/GR157368 Complex. The synthesis of GR157368 was an attempt to define the "minimal core" which would be required for binding and would act as a scaffold for further interactions with the protein (*30*); see Figure 2. The crystal structure of thrombin/GR157368 was determined to identify the precise position of this minimal core template (complex A). This structure showed the acylation of the active site Ser195 of thrombin by GR157368, resulting in a ring-opened conformation of the *trans*-lactone template and a covalent bond between the carbonyl carbon of the lactone and the Oγ of the serine side chain (Figure 3). The carbonyl oxygen of the ring-opened lactone resides in the oxyanion hole making hydrogen bonds to the backbone NH atoms of Gly193 (2.8 Å) and Ser195 (2.8 Å). A water-mediated interaction is observed between the inhibitor hydroxyl and NZ of Lys60f from the 60-insertion loop; a similar interaction was also observed in the structures of thrombin complexed to GR133686 and GR133487 (*24*). Interactions in the S1 specificity pocket are dominated by the basic nature of the



(b)

Complex	Chemical series	IC_{50} (nM)	$k_{obs/I}$ ($\times 10^6 M^{-1} s^{-1}$)
A	cyclopentane	130	0.004
B	6-series indane	9	0.265
C	5-series indane	18	0.096
D	5-series indane	8	0.420

GR157368

GR166081

GR167088

GR179849

FIGURE 2: (a) Chemical structure of the lead compound GR133686 drawn with respect to its binding mode in the thrombin active site, as reported by Weir et al. (*24*). The enone side chain is bound in the S1 pocket, and the S2 pocket is partially occupied by a methyl group. The jagged line indicates the bond that is broken on acylation of the active site Ser195. (b) Chemical structures of the inhibitors shown together with inhibition data against thrombin. Apparent second-order association rate constants ($k_{obs/I}$) were determined essentially as described by Weir et al. (*24*). For comparison, the apparent second-order rate constant for PPACK is 12×10^6 M^{-1} s^{-1} as reported by Kettner and Shaw (*39*).

alkylamidine substituent, and as expected the amidine forms a salt bridge with the carboxylate side chain of Asp189 which is located at the base of the pocket; distances between amidine nitrogens and carboxylate oxygens are 2.9−3.0 Å (Figure 4a). It is notable that the bound conformation of the alkylamidine side chain contains a gauche torsion angle.

It is clear from complex A that the only binding pocket of thrombin that is occupied is S1, and in order to access the S2 and S3 pockets, further modifications to the *trans*-lactone template would be required. To enable this, an aromatic ring was fused onto the *trans*-lactone template which could then be substituted at different positions to provide suitable access to the S3 pocket (*30*). This modification was consistent with the binding mode of GR133686 to thrombin which defined the position of the core euphane triterpene of GR133686 in the active site and indicated how the *trans*-lactone template could be changed to increase bulk but avoid steric clashes (*24*). On the basis of this structural information, two series of indane *trans*-lactone compounds were synthesized by substituting at the 5 or 6 position of the phenyl ring; three of these compounds (GR166081, GR167088, GR179849) and their inhibitory activity against thrombin are shown in Figure 2b. Structure−activity relationships (SAR) of the two series indicated that the "6-series" inhibitors were consistently more active against thrombin and also more selective for thrombin over trypsin (*30*). To explore the molecular basis of this difference in potency and selectivity of the two series of inhibitors, crystal structures of thrombin complexed to representatives from both the 5- and 6-series were determined.

Thrombin/GR166081 Complex. The structure of thrombin complexed to the 6-series inhibitor GR166081 (IC_{50} = 9nM) was solved and analyzed (complex B) (Figure 2). The electron density in the difference Fourier map of the inhibitor

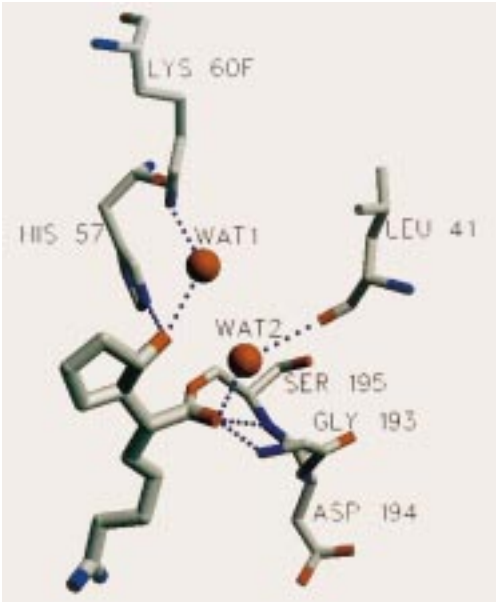


FIGURE 3: Key interactions between GR157368 and thrombin around the oxyanion hole. Dotted lines indicate hydrogen bonds, the inhibitor is represented by thicker bonds, and the red spheres are water molecules. A covalent bond is drawn between the Ser195 Oγ and the inhibitor carbonyl which represents acylation.

molecule was sufficient to define its conformation, and it was clear that the *trans*-lactone had been opened and the resulting carbocyclic template bound in a manner similar to that seen for complex A (Figure 1). The lactone ring has been opened by Ser195, and the resulting ester carbonyl oxygen atom is involved in hydrogen-bonding interactions with the oxyanion hole (Figure 4b). Also, the water-mediated interaction between the inhibitor hydroxyl and the Lys60f side chain is conserved. Although the two complexes A and B exhibit grossly similar inhibitor binding modes, there is a difference in the orientation of the five-membered carbocyclic ring; in complex B the ring is positioned closer to the S2 pocket than in complex A. This may be a consequence of the additional hydrophobic phenyl ring of GR166081, which is fused onto the carbocyclic template, partially binding in the S2 pocket and so forcing the inhibitor to adopt a different orientation compared to that of GR157368. Although the second phenyl ring of GR166081 is located at the mouth of the S3 pocket, where it makes van der Waals contact with the backbone atoms of Gly216, the hydroxyl group on the ring makes no specific interactions with the protein. Despite the difference in the position of the carbocyclic template in complexes A and B, the interactions in the S1 pocket are very similar; the alkylamidine of GR166081 also makes a salt bridge interaction with Asp189 and a hydrogen bond to Gly219. Furthermore, the conformation of the alkylamidine side chain, which again has a gauche bond in the chain, is similar in both structures.

It was apparent from the structure of complex B that access to the S3 pocket could be obtained by substitution at the meta position on the aryl ring which is located at the entrance to the pocket. One such compound which contains a meta-substituted aryl ring but from the 5-series is GR167088 (IC_{50} = 18nM); see Figure 2. Molecular modeling studies based on complex B suggested that the methoxy substituent of GR167088 would bind in the S3 pocket. To confirm this prediction, and also to define the binding mode of a 5-series inhibitor, the thrombin/GR167088 structure was determined (complex C).

Thrombin/GR167088 Complex. The indane carbocyclic template in complex C displayed a binding mode similar to that observed in the previous two complexes due to the acylation of Ser195 (Figure 4c). In complex C, however, the phenyl ring of the indane did not occupy the S2 pocket to any significant degree although its position blocked access to the pocket. Moreover, the methoxy-substituted aryl ring was located at the entrance to the S3 pocket with the methoxymethyl group bound further into the pocket making hydrophobic interactions with the side chains of Trp215, Leu99, and Ile174. The packing of the methoxymethyl in the S3 pocket is not optimal as the closest contact distance between the methyl carbon and any protein atom in this pocket is 4 Å. This binding mode was largely as predicted, based on visual inspection of the previous complexes, apart from the significant lack of binding in the S2 pocket. Analysis of the thrombin S3 pocket in complex C indicated that a substituent more bulky than a methoxy could be accommodated. The hydrophobic nature of the pocket suggested neutral substituents would be preferred; however, the hydroxyl of Tyr60A provided some potential for hydrogen bonding. The 5-series inhibitor GR179849, in which a diethylcarboxamide group replaced the methoxy group of GR166088, was synthesized in an attempt to exploit the hydrophobic nature of the S3 pocket. To define the binding mode of GR179849 and ascertain whether the predictions were correct, the crystal structure of the complex was determined (complex D).

Thrombin/GR179849 Complex. Difference electron density for the inhibitor in complex D was largely well defined, except for some local disorder in the diethylcarboxamide region, and allowed a confident interpretation to be made (Figure 1). The interactions observed in complex D are very similar to those in complex C; the inhibitor molecules overlap extremely well with significant differences localized to the S3 pocket as expected (Figure 4d). In complex D the diethylcarboxamide occupies the S3 pocket with one of the ethyl groups making good hydrophobic interactions with the side chain of Ile174 (closest contact distance is 3.7 Å) and Trp215 (closest contact distances are between 3.3 and 3.4 Å). The other ethyl group appears to be less well ordered and is located almost outside the S3 pocket but may pack against the side chain of Leu99 (distance between CG and the ethyl carbon atom is 3.9 Å). Although the distance between the carbonyl oxygen and the OH of Tyr60a is consistent with a hydrogen bond (3.1 Å), it is likely to be a weak interaction due to the accessibility of the tyrosine hydroxyl to solvent. Despite the apparent disorder of the diethylcarboxamide group, it is clear from the electron density that it is located in the S3 pocket. Modeling studies had also suggested an alternative binding mode for GR179849 which would result in an interaction between the carboxamide carbonyl and the backbone NH of Gly219; however, this was not observed in the crystal structure.

DISCUSSION

The binding modes displayed by this novel series of active site-directed thrombin inhibitors clearly show a large degree of conservation which is largely a consequence of the generic

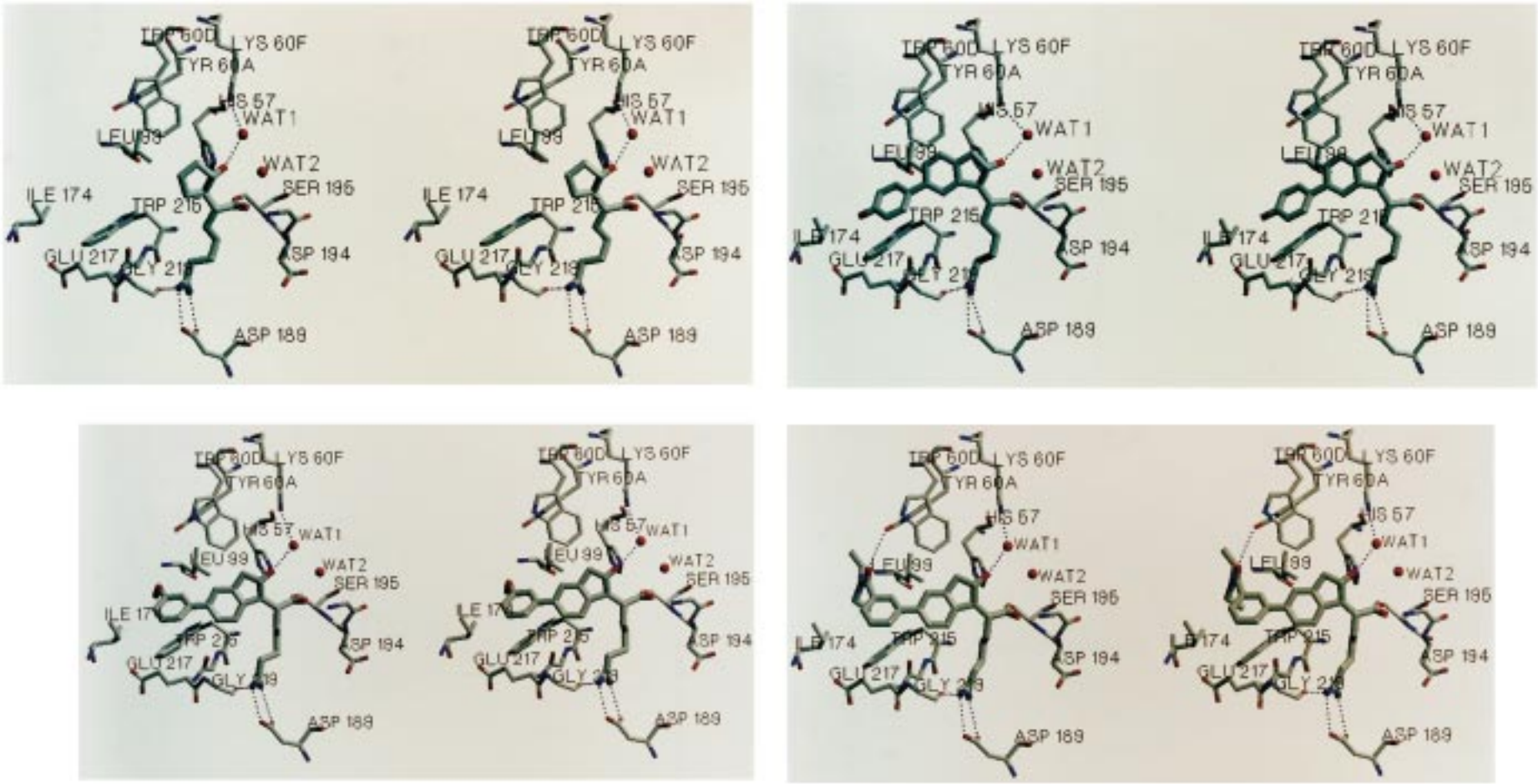


FIGURE 4: Stereoviews of the thrombin active site with key protein/inhibitor interactions highlighted for four complexes: (a, top left) GR157368, (b, top right) GR166081, (c, bottom left) GR167088, and (d, bottom right) GR179849. In each view hydrogen bonds are represented by dotted lines, and the inhibitor is represented by thicker bonds.

trans-lactone-based mechanism of inhibition. In each case the potential interactions within the S1 pocket and the oxyanion hole are constrained due to the acylation of the active site serine, tethering the inhibitor to the enzyme. It was due to this conservation of binding that we were able to utilize the core *trans*-lactone template as a scaffold and model additional functionalities to interact with the other protein pockets. These crystal structures are the first to describe binding modes of a series of nonpeptidic synthetic inhibitors that form a covalent adduct with thrombin. There has been, however, a previous report of a series of peptidic boronic acid inhibitors of thrombin whose binding modes were also studied by X-ray crystallography (*32*). In these crystal structures the formation of a covalent adduct is observed between the Ser195 hydroxyl and the boron atom of the inhibitor. Interestingly, there is also a conserved water-mediated interaction (via WAT2 as labeled in their study) between an inhibitor hydroxyl and Lys60f from the 60-insertion loop, a recurring feature seen in the *trans*-lactone complexes. An analogous solvent-mediated interaction with Lys60f is also observed on binding of NAPAP and MD-805 to thrombin (*13*). Future inhibitor design experiments could incorporate a direct interaction with Lys60f that would be expected to impart selectivity as this residue is unique to thrombin.

Another conserved solvent interaction identified in the peptidic boronate series and in the *trans*-lactone series is a water molecule located close to the oxyanion hole. In the boronate complexes this water (WAT1) hydrogen bonds to an inhibitor hydroxyl as well as to the backbone NH of Gly193 and the CO of Leu41. A water molecule (WAT2) in a similar location and with equivalent interactions with the protein, and which also hydrogen bonds to the inhibitor carbonyl in the oxyanion hole, is observed in all four of the resulting carbocyclic complexes described above (Figure 3). In fact, a water molecule in this position is conserved in many serine protease/ligand complexes and has been implicated by some workers in the deacylation mechanism of peptide hydrolysis (*33*). However, recent studies based on a stable acyl-enzyme complex of porcine pancreatic elastase have identified an alternative and more likely position for the deacylating water molecule (*34*). This position for the water molecule is also consistent with the proposed mechanism for the relactonization of the 5,5-*trans*-lactone-containing inhibitor GR133487 (*24*).

A good example of optimal recognition by a thrombin inhibitor is highlighted by the crystal structure of the thrombin/PPACK complex (*35*). The functional groups within PPACK mimic the natural substrate specificity of thrombin, and this excellent molecular complementarity is reflected in the potency of the inhibitor (Figure 2). It is therefore instructive to compare the key features of the thrombin/PPACK complex with the binding mode exhibited by GR179849, a representative of the *trans*-lactone series of inhibitors. Occupation of the S1 pocket by the alkylamidine substituent of GR179849 results in good electrostatic interactions between the basic amidine nitrogens and the carboxylate of Asp189 (Figure 5). This salt bridge is also observed in PPACK, and in addition, the Nϵ atom of the arginylguanidinium group of PPACK hydrogen bonds to the carbonyl of Gly219 via a bridging water molecule; this nitrogen is absent in the alkylamidine side chain. The

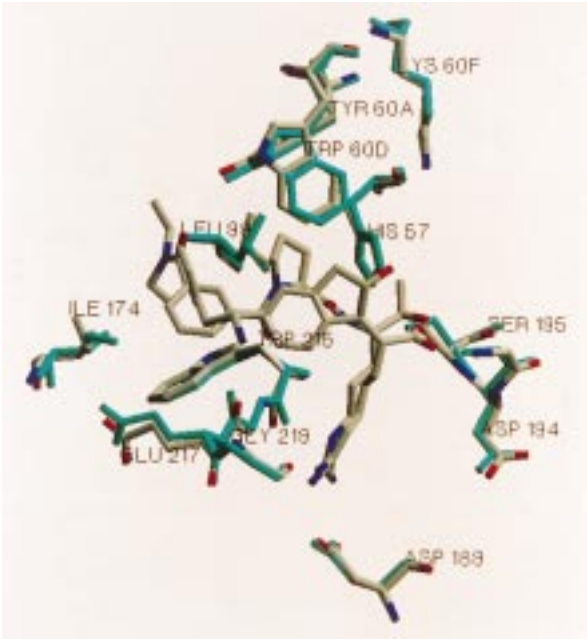


FIGURE 5: Superposition of the thrombin/PPACK structure with complex D (GR179849) showing the bound inhibitor molecules at the active site. The protein residues from the PPACK structure are colored cyan. Apart from the side chain of Ile174, the conformations of the residues around the active site from the two complexes are very similar.

conformation of the P1 arginyl side chain of PPACK is extended; in comparison there is a gauche torsion angle in the alkylamidine side chain; similar geometry is observed in the other three *trans*-lactone complexes. As the lengths of these P1 side chains are identical and they both position the terminal nitrogen atoms in a very similar location at the base of the pocket, the different conformations may, in part, be due to the alkylation of His57 by PPACK as this causes the Cα of the arginine residue to be located in a different position compared to the analogous carbon atom in the *trans*-lactone inhibitor (Figure 5). Thus, the direction of entry into the S1 pocket of the two P1 side chains is significantly different. Also, the substitution of the sp^2-hybridized Nϵ atom of the arginylguanidinium by an sp^3 carbon in the alkylamidine may also be relevant as the minimum energy conformations for these side chains are likely to differ. Preliminary conformational analysis of the alkylamidine side chain performed in the enzyme pocket suggests that an extended conformation is not possible due to steric clashes and also that the observed conformation is actually a low-energy conformer (data not shown). One of the peptidic boronate inhibitors, from the series mentioned above, also contains an alkylamidine P1 group (*32*). The P1 side chain conformation observed in this structure (PDB code: 1LHE) is very similar to the alkylamidine conformation of GR157368 observed in complex A, i.e., gauche.

Comparison of GR179849 in complex D with the PPACK complex identifies a key difference in occupation of the S2 pocket (Figure 5). For PPACK a proline residue packs into the hydrophobic S2 pocket and appears to be the optimal shape and size. However, in complex D there is no significant binding in the S2 pocket; the phenyl ring of the indane is located at the mouth of the pocket. In complex B (GR166081), however, there appears to be more occupation of the S2 pocket as the hydrophobic phenyl ring of the indane is

located in a similar position to the proline side chain of PPACK. This difference in S2 binding between a 5-series (GR179849) and a 6-series (GR166081) inhibitor is reflected in SAR, which suggest that not only are the 6-series inhibitors more potent but they are more selective for thrombin against related serine proteases such as trypsin (*30*). Invoking the thrombin S2 pocket to help determine selectivity is clearly rationalized as the pocket is largely formed by the 60-insertion loop, an insertion loop unique to thrombin. Several other structural studies of thrombin inhibitors have implicated the S2 pocket as key for obtaining selectivity for thrombin over related serine proteases; these include series based on the 5-amidinoindole (*18*), bi- and tricyclic heterocycles (*19*), and lactam sulfonamide templates (*36*).

The D-Phe residue of PPACK is able to bind in the S3 pocket due to its unnatural stereochemistry; the benzyl ring makes an aryl−aryl (edge-on) interaction with the side chain of Trp215 and also good van der Waals contacts with Leu99 and Ile174. From the *trans*-lactone series only GR179849 (complex D) is able to occupy the S3 pocket of thrombin; the diethylcarboxamide buries one of its ethyl groups into this hydrophobic pocket. A hydrogen-bonding interaction between the carboxamide carbonyl and the OH of Tyr60a (3.1 Å) may also provide additional binding energy for this compound. A similar hydrogen-bonding interaction was observed on binding a piperonyl substituent in the S3 pocket from a series of reversible thrombin inhibitors which contained bi- and tricyclic core templates (*19*). Many thrombin inhibitors, in particular those more peptidic in nature, attempt to utilize the potential hydrogen bonds to the backbone of Gly216 as these are implicated in thrombin/substrate interactions and so may provide significant binding energy (*20*, *22*, *37*). Indeed, PPACK hydrogen bonds to both the carbonyl and the backbone NH of Gly216, resulting in an antiparallel β-sheet interaction. Although no attempt was made to incorporate this interaction into the indane *trans*-lactone template, there may be a weak hydrogen bond interaction between a CH of the phenyl ring of the indane, where the hydrogen will have a partial positive charge, and the backbone carbonyl of Gly216. This interaction is only possible for the 5-series compounds GR167088 and GR179849 where the relevant distance between the CH carbon atom and the oxygen of the carbonyl is 3.2 Å. There are several examples of these types of hydrogen-bonding interactions between acidic hydrogens and carbonyl oxygen atoms that have been previously reported in protein crystal structures (*38*).

Potency of the *trans*-lactone series was assessed by measuring the apparent association rate constant ($k_{obs/I}$) in addition to the IC_{50} values (*24*). From the limited SAR reported in Figure 2, it is clear that we were able to improve binding to thrombin by the structure-based modifications described in this study. For example, the improvement in binding by better occupation of the S3 pocket is reflected by GR167088 and GR179849, thus comparing a methoxy group with a diethylcarboxamide, respectively. The structural data also helped to rationalize the differences in potency and selectivity for thrombin between the 5- and 6-series of the indane *trans*-lactone inhibitors; better occupation of the S2 pocket was evident for GR166081. Furthermore, the crystal structures described in this study were able to identify the key elements of recognition by the *trans*-lactone inhibitors of thrombin. Finally, it is envisaged that the novel reactive *trans*-lactone moiety central to these inhibitors could also be utilized for inhibition of other therapeutically important serine proteases.

ACKNOWLEDGMENT

We thank Dr. Richard Dennis for the mass spectrometric analysis, Dr. Sue Bethell for advice during purification, and Dr. Mike Hann, Dr. Rob Cooke, and Dr. Nigel Ramsden for critically reading the manuscript. We also thank Dr. Darko Butina for contributions to the molecular modeling work and Dr. Chris Mooney for discussions on kinetic parameters. We also acknowledge useful discussions with Dr. Nigel Watson and Dr. Harry Finch.

REFERENCES

1. Davie, E. W., Fujikawa, K., and Kiesel, W. (1991) *Biochemistry 30*, 10363−10370.
2. Furie, B., and Furie, B. C. (1988) *Cell 53*, 505−518.
3. Vu, T.-K. H., Hung, D. H., Wheaton, V. I., and Coughlin, S. R. (1991) *Cell 64*, 1057−1068.
4. Stone, S. R., and Hofsteenge, J. (1986) *Biochemistry 25*, 4622−4628.
5. van de Locht, A., Lamba, D., Bauer, M., Huber, R., Friedrich, T., Kroger, B., Hoffken, W., and Bode, W. (1995) *EMBO J. 14*, 5149−5157.
6. Harker, L. A., Hanson, S. R., and Kelly, A. B. (1997) *Thromb. Haemostasis 78*, 736−741.
7. Berliner, L. J. (1992) *Thrombin: Structure and Function* (Berliner, L. J., Ed.) Plenum Press, New York.
8. Stubbs, M. T., and Bode, W. (1993) *Thromb. Res. 69*, 1−58.
9. Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S. R., and Hofsteenge, J. (1989) *EMBO J. 8*, 3467−3475.
10. Bode, W., Brandstetter, H., Mather, T., and Stubbs, M. (1997) *Thromb. Haemostasis 78*, 501−511.
11. Powers, J. C., and Kam, C.-M. (1992) *Thrombin: Structure and Function* (Berliner, L. J., Ed.) pp 117−158, Plenum Press, New York.
12. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
13. Banner, D. W., and Hadvary, P. (1991) *J. Biol. Chem. 266*, 20085−20093.
14. Rydel, T. J., Ravichandran, K. G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C., and Fenton, J. W., II (1990) *Science 249*, 277−280.
15. Qiu, X., Padmanabhan, K. P., Carperos, V. E., Tulinsky, A., Kline, T., Maraganore, J. M., and Fenton, J. W., II (1992) *Biochemistry 31*, 11689−11697.
16. Martin, P. D., Robertson, W., Turk, D., Huber, R., Bode, W., and Edwards, B. F. P. (1992) *J. Biol. Chem. 267*, 7911−7920.
17. Martin, P. D., Malkowski, M. G., DiMaio, J., Konishi, Y., Ni, F., and Edwards, B. F. P. (1996) *Biochemistry 35*, 13030−13039.
18. Chirgadze, N. Y., Sall, D. J., Klimkowski, V. J., Clawson, D. K., Briggs, S. L., Hermann, R., Smith, G. F., Gifford-Moore, D. S., and Wery, J.-P. (1997) *Protein Sci. 6*, 1412−1417.
19. Obst, U., Banner, D. W., Weber, L., and Diederich, F. (1997) *Chem. Biol. 4*, 287−295.
20. Sanderson, P. E. J., Dyer, D. L., Naylor-Olsen, A. M., Vacca, J. P., Gardell, S. J., Lewis, S. D., Lucas, B. J., Jr., Lyle, E. A., Lynch, J. J., Jr., and Mulichak, A. M. (1997) *Bioorg. Med. Chem. Lett. 7*, 1497−1500.
21. Engh, R. A., Brandstetter, H., Sucher, G., Eichinger, A., Baumann, U., Bode, W., Huber, R., Poll, T., Rudolph, R., and von der Saal, W. (1996) *Structure 4*, 1353−1362.
22. Das, J., and Kimball, S. D. (1995) *Bioorg. Med. Chem. 3*, 999−1007.

23. O'Neill, M. J., Lewis, J. A., Noble, H. M., Holland, S., Mansat, C., Farthing, J. E., Foster, G., Noble, D., Lane, S. J., Sidebottom, P. J., Lynn, S., Hayes, M. V., and Dix, C. J. (1998) *J. Nat. Prod*. (in press).
24. Weir, M. P., Bethell, S. S., Cleasby, A., Campbell, C. J., Dennis, R. J., Dix, C. J., Finch, H., Jhoti, H., Mooney, C. J., Patel, S., Tang, C.-M., Ward, M., Wonacott, A. J., and Wharton, C. W. (1998) *Biochemistry 37*, 6645−6657.
25. Ngai, P. K., and Chang, J.-Y. (1991) *Biochem. J. 280*, 805−808.
26. Lottenberg, R., sen, U., Jackson, C. M., and Coleman, P. L. (1981) *Methods Enzymol. 80*, 341−361.
27. Skrzypczak-Jankun, E., Carperos, V. E., Ravichandran, K. G., and Tulinsky A. (1991) *J. Mol. Biol. 221*, 1379−1393.
28. Messerschmidt, A., and Pflugrath, J. W. (1987) *J. Appl. Crystallogr. 20*, 306−315.
29. CCP4: Collaborative Computor Project No. 4. The CCP4 Suite: Programs for Protein Crystallography (1994) *Acta Crystallogr. D50*, 760−763.
30. Pass, M., Bolton, R. E., Coote, S. J., Finch, H., Hindley, S., Lowdon, A., McLaren, J., Owen, M., Pegg, N. A., Mooney, C. J., Tang, C.-M., Parry, S., and Patel, C. (1998) *Bioorg. Med. Chem. Lett.* (in press).
31. Brunger, A. T., Kuriyan, J., and Karplus, M. (1987) *Science 235*, 458−460.
32. Weber, P. C., Lee, S.-L., Lewandowski, F. A., Schadt, M. C., Chang, C.-H., and Kettner, C. A. (1995) *Biochemistry 34*, 3750−3757.
33. Perona, J. J., Craik, C. S., and Fletterick, R. J. (1993) *Science 261*, 620.
34. Wilmouth, R. C., Clifton, I. J., Robinson, C. V., Roach, P. L., Aplin, R. T., Westwood, N. J., Hajdu, J., and Schofield, C. J. (1997) *Nat. Struct. Biol. 4*, 456−462.
35. Bode, W., Turk, D., and Karshikov, A. (1992) *Protein Sci. 1*, 426−471.
36. Semple, J. E., Rowley, D. C., Brunck, T. K., Ha-Uong, T., Minami, N. K., Owens, T. D., Tamura, S. Y., Goldman, E. A., Siev, D. V., Ardecky, R. J., Carpenter, S. H., Ge, Y., Richard, B. M., Nolan, T. G., Hakanson, K., Tulinsky, A., Nutt, R. F., and Ripka, W. C. (1996) *J. Med. Chem. 39*, 4531−4536.
37. Tamura, S. Y., Goldman, E. A., Brunck, T. K., Ripka, W. C., and Semple J. E. (1997) *Bioorg. Med. Chem. Lett. 7*, 331−336.
38. Derewenda, Z. S., Lee, L., and Derewenda, U. (1995) *J. Mol. Biol. 252*, 248−262.
39. Kettner, C., and Shaw, E. (1979) *Thromb. Res. 14*, 969−973.

BI9830359