Molecular recognition at the thrombin active site: structure-based design and synthesis of potent and selective thrombin inhibitors and the X-ray crystal structures of two thrombin–inhibitor complexes

Ulrike Obst[1], David W Banner[2], Lutz Weber[2] and François Diederich[1]

Background: The serine protease thrombin is central in the processes of hemostasis and thrombosis. To be useful, thrombin inhibitors should combine potency towards thrombin with selectivity towards other related enzymes such as trypsin. We previously reported the structure-based design of thrombin inhibitors with rigid, bicyclic core structures. These compounds were highly active towards thrombin, but showed only modest selectivity.

Results: Here, we describe the rational design of selective thrombin inhibitors starting from the X-ray crystal structure of the complex between the previously generated lead molecule and thrombin. The lead molecule bound with a K_i value of 90 nM and a selectivity of 7.8 for thrombin over trypsin. Our design led to inhibitors with improved activity and greatly enhanced selectivity. The binding mode for two of the new inhibitors was determined by X-ray crystallography of their complexes with thrombin. The results confirmed the structures predicted by molecular modeling and, together with the binding assays, provided profound insight into molecular recognition phenomena at the thrombin active site.

Conclusions: A novel class of nonpeptidic, selective thrombin inhibitors has resulted from structure-based design and subsequent improvement of the initial lead molecule. These compounds, which are preorganized for binding to thrombin through a rigid, bicyclic or tricyclic central core, could aid in the development of new antithrombotic drugs. Correlative binding and X-ray structural studies within a series of related, highly preorganized inhibitors, which all prefer similar modes of association to thrombin, generate detailed information on the strength of individual intermolecular bonding interactions and their contribution to the overall free energy of complexation.

Addresses: [1]Laboratorium für Organische Chemie, ETH-Zentrum, Universitätstrasse 16, 8092 Zürich, Switzerland and [2]Pharma Division, Preclinical Research, F Hoffmann-La Roche Ltd, 4002 Basel, Switzerland.

Correspondence: François Diederich
E-mail: diederich@org.chem.ethz.ch

Key words: molecular recognition, protein X-ray crystallography, structure-based inhibitor design, thrombin inhibitors

Received: **21 February 1997**
Accepted: **24 March 1997**

Chemistry & Biology April 1997, 4:287–295
http://biomednet.com/elecref/1074552100400287

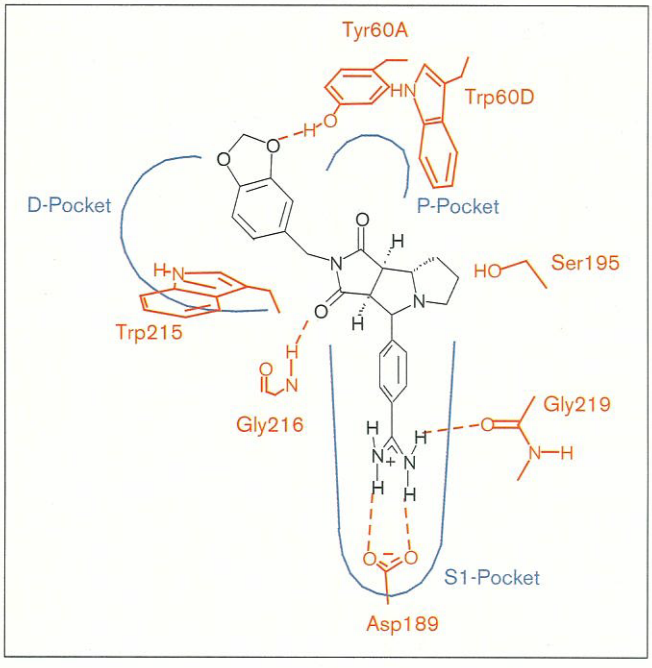
Introduction

Thrombin is a trypsin-like serine protease that is central in the processes of hemostasis and thrombosis. In the blood coagulation cascade, it cleaves the protein fibrinogen at specific arginine residues to give polymerizable fibrin, which is a major constituent of blood clots. Moreover, thrombin is the main activator of platelet aggregation and other coagulation factors [1]. Defects in the delicate balance between coagulation factors and their endogenous inhibitors can lead to serious complications such as the formation of thrombi in blood vessels. Thrombin is therefore an important pharmaceutical target for the treatment and prevention of thrombotic diseases. The search for small, potentially orally bioavailable synthetic inhibitors of thrombin is especially intense [2,3]. X-ray crystal structures of thrombin–inhibitor complexes [4,5] show that the enzyme is a rather rigid protein with well-defined binding pockets in the active site. These properties make thrombin a particularly well-suited target for structure-based inhibitor design [6–8].

We previously described [9] the rational design of novel, nonpeptidic thrombin inhibitors with rigid, bicyclic core structures. Although these compounds were highly active for thrombin, their selectivity for thrombin compared with the structurally related digestive enzyme trypsin was not satisfactory. The most active compound in this series, *rac*-**1** (*rac* = racemic; Fig. 1), bound to thrombin with $K_i = 90$ nM and displayed a selectivity [K_i(trypsin)/K_i(thrombin)] of 7.8. Other serine proteases in the blood coagulation cascade, such as factors VIIa and Xa, are also structurally related to thrombin. Thus, thrombin inhibitors for potential drug development need to discriminate against these proteases.

In this work, we present the design and synthesis of novel, nonpeptidic, selective thrombin inhibitors and we confirm their binding mode by X-ray crystallography on two thrombin–inhibitor complexes. An elegant optical resolution of the new inhibitors is described. Binding affinity

Figure 1



Schematic representation of the complex formed by thrombin and *ent*-**1**, according to the X-ray crystal structure [9].

data correlated with X-ray structural information provide a detailed insight into the nature and strength of weak intermolecular bonding interactions at the thrombin active site.

Results and discussion

Inhibitor design

To increase the selectivity of new inhibitors for thrombin over trypsin, the X-ray crystal structure of the complex formed between thrombin and the lead molecule **1** (Fig. 1) [9] was carefully examined. Only one enantiomer of the racemic mixture of compound **1** was found in the active site of thrombin, as expected from molecular modeling experiments [10]. This enantiomer was named *ent*-**1** (*ent* = enantiomeric).

The amidinium sidechain of *ent*-**1** binds well in the recognition pocket S1 of the enzyme. It interacts with the carboxylate group of Asp189 at the bottom of the pocket and forms additional hydrogen bonds to the carbonyl group of Gly219 and to a water molecule (not shown in Fig. 1). The aromatic part of the piperonyl (benzo[1,3]dioxol-5-ylmethyl) group of *ent*-**1** is located in the large hydrophobic D-pocket of thrombin (called D because it is 'distal' from the catalytic Ser195 [5]) and is involved in a CH···π interaction with Trp215. One oxygen atom of the piperonyl group forms a hydrogen bond with Tyr60A. Another hydrogen bond exists between the NH of Gly216 and an imide C=O group of *ent*-**1**. The other, 'upper' C=O group of the inhibitor points into the small, hydrophobic P-pocket ('proximal') of thrombin. The transfer of this C=O group from aqueous solution to the hydrophobic pocket requires desolvation energy and is therefore unfavorable. The P-pocket is formed by an insertion loop in the structure of the enzyme and is lacking in other, related serine proteases such as trypsin. This P-pocket is unique to thrombin and its occupancy by a stereoelectronically complementary group provides a significant contribution to the narrow substrate specificity of the enzyme. Therefore, by removing the 'upper' C=O group in *ent*-**1**, which lacks this complementarity, or by replacing it with hydrophobic residues, we expected binding selectivity to be increased.

Synthesis

The key step in the synthesis of the starting materials *rac*-**2** and *rac*-**3** involved the formation of the central bicyclic skeleton by an azomethine ylide 1,3-dipolar cycloaddition [11]. The synthesis was carried out as previously described [9]. The conversion of one C=O into one CH_2 group in *rac*-**2** was achieved by treatment with $Li[Et_3BH]$ in tetrahydrofuran (THF) followed by $Na[BH_3CN]$ in trifluoroacetic acid (TFA) (Fig. 2). This reaction sequence yielded both regioisomeric lactams *rac*-**4** and *rac*-**5**, which were separated and converted into the corresponding nitriles *rac*-**6** and *rac*-**7**, respectively, using CuCN in refluxing dimethylformamide (DMF). The synthesis of the amidinium salts *rac*-**8** and *rac*-**9** was completed by a Pinner reaction [12].

For the replacement of the upper C=O group by a hydrophobic residue, *rac*-**2** was reduced with $Li[Et_3BH]$ to yield a mixture of two regioisomeric hydroxylactams (Fig. 2). After separation, reaction with toluene-4-sulfinic acid afforded the *exo*-sulfone *rac*-**10** [13]. Nucleophilic displacement by organometallic reagents prepared from the corresponding Grignard reagents and zinc chloride afforded the *exo*-alkyl-substituted or cycloalkyl-substituted lactams *rac*-**11a,b,d,e** [13], which were converted via the nitriles *rac*-**12a,b,d,e** into the amidinium salts *rac*-**13a,b,d,e**.

Attempts to find a more direct way of replacing one C=O group of *rac*-**2** by reaction with MeMgCl followed by reduction with $Na[BH_3CN]$ gave the *endo*-methylated lactam *rac*-**14** as the only product, which was also converted via nitrile *rac*-**15** into the amidinium salt *rac*-**16**.

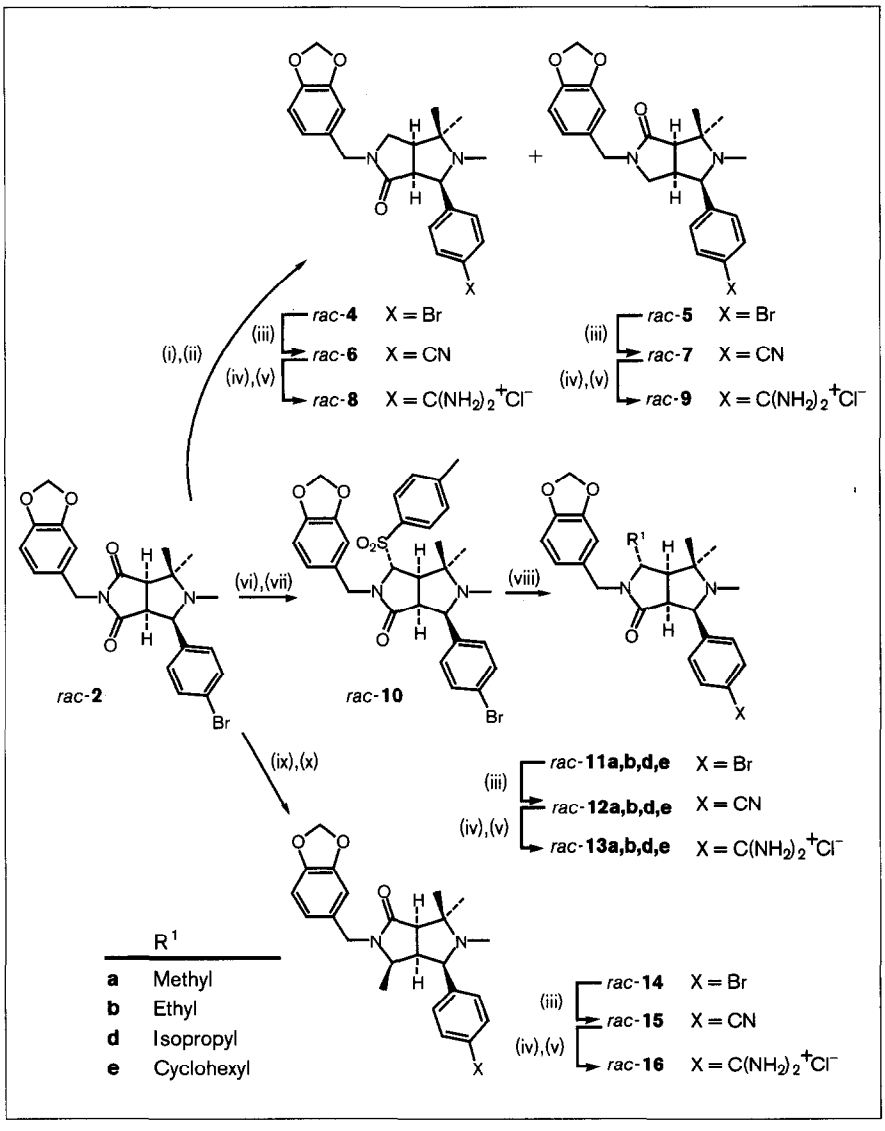
The same sequence of synthetic steps was used for the transformation of imide *rac*-**3**, with a tricyclic core structure, into the *exo*-sulfone *rac*-**17** and, via *rac*-**18b–f** and *rac*-**19b–f**, into the amidinium salts *rac*-**20b–f** (Fig. 3).

Biological results

With their rigid central bicyclic or tricyclic cores, all inhibitors discussed in this study have highly constrained conformations and it is therefore reasonable to assume

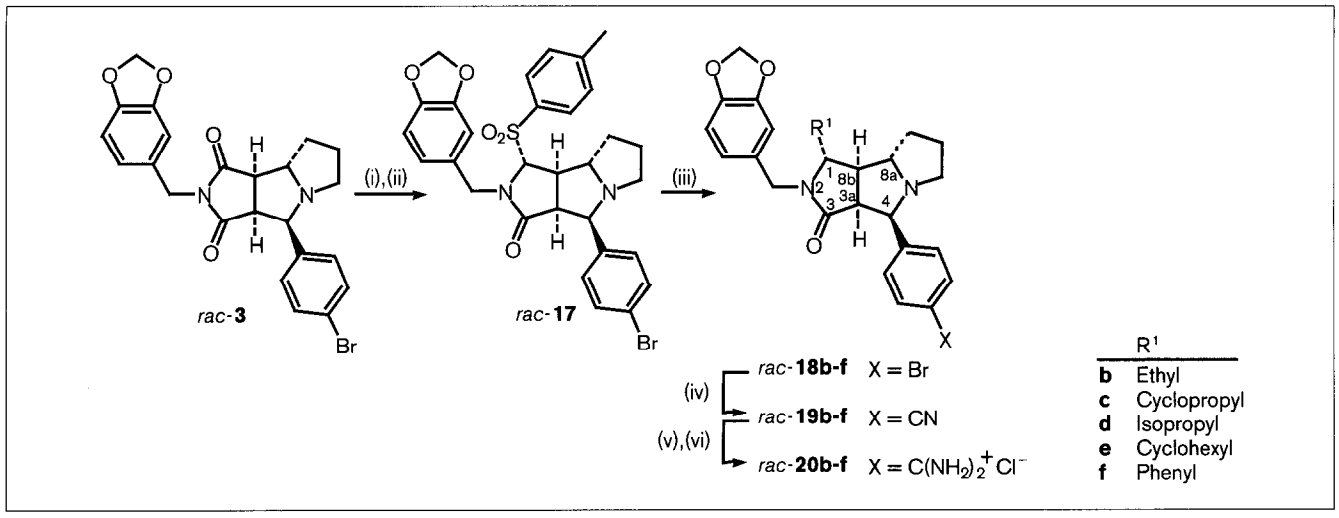
Figure 2

Synthesis of the racemic inhibitors **8**, **9**, **13a,b,d,e** and **16**. (i) $Li[Et_3BH]$, tetrahydrofuran (THF), −78→0°C, 1 h. (ii) $Na[BH_3CN]$, trifluoroacetic acid (TFA), room temperature (r.t.), 5 h, 49% (*rac*-**4**) and 31% (*rac*-**5**). (iii) CuCN, DMF, reflux, 17–30 h, 62% (*rac*-**6**), 65% (*rac*-**7**), 62% (*rac*-**12a**), 40% (*rac*-**12b**), 55% (*rac*-**12d**), 62% (*rac*-**12e**), 65% (*rac*-**15**). (iv) CH_3OH, HCl (gas), $CHCl_3$, 4°C, 24 h. (v) NH_3, CH_3OH, 65°C, 3.5 h, 84% (*rac*-**8**), 82% (*rac*-**9**), 90% (*rac*-**13a**), 90% (*rac*-**13b**), 66% (*rac*-**13d**), 74% (*rac*-**13e**), 83% (*rac*-**16**). (vi) $Li[Et_3BH]$, THF, −78→0°C, 2 h, 67%. (vii) Toluene-4-sulfinic acid, $CaCl_2$, CH_2Cl_2, r.t., 19 h, 74%. (viii) $ZnCl_2$, RMgCl or RMgBr, CH_2Cl_2, 0°C→r.t., 13–36 h, 75% (*rac*-**11a**), 76% (*rac*-**11b**), 35% (*rac*-**11d**), 31% (*rac*-**11e**). (ix) MeMgCl, THF, r.t., 13 h. (x) $Na[BH_3CN]$, MeOH, TFA, r.t., 2 h, 94%.



that they each bind to thrombin with an identical binding mode. Using this hypothesis, which is further validated by the X-ray crystallographic findings described below, we could evaluate the contributions made by single substituents on the central inhibitor core to the measured free energy of binding. Table 1 shows the binding affinities [14,15] of the bicyclic inhibitors with various residues at C(4). A comparison of the affinities of imide *rac*-**21**, which was prepared as for *rac*-**1** [9], and lactam *rac*-**8** towards thrombin indicated, quite surprisingly, that the imide bound by 1.5 kcal mol^{-1} better than the lactam. Apparently, it is more favorable to have a C=O group of poor electronic complementarity filling the hydrophobic P-pocket of the enzyme and pointing onto the surfaces of the aromatic rings of Tyr60A and Trp60D than to leave this pocket empty. When the inhibitor (*rac*-**13a**) had an *exo*-methyl group at C(4), which should bind into the P-pocket, both the affinity ($\Delta\Delta G_{8\rightarrow 13a} = -0.7$ kcal mol^{-1}) and the selectivity increased compared to lactam *rac*-**8**. Nevertheless, the methylated lactam *rac*-**13a** remained less potent than imide *rac*-**21**. A large improvement in both the affinity and the selectivity, compared to *rac*-**8**, resulted from the introduction of an ethyl group (*rac*-**13b**, $\Delta\Delta G_{8\rightarrow 13b} = -2.5$ kcal mol^{-1}), and an isopropyl group improved parameters still further (*rac*-**13d**, $\Delta\Delta G_{8\rightarrow 13d} = -3.2$ kcal mol^{-1}). With a K_i value of 30 nM and a 223-fold higher affinity for thrombin over trypsin, *rac*-**13d** was the most potent and most selective among the inhibitors with a bicyclic core. A cyclohexyl substituent is too large for optimal binding in the P-pocket and both the affinity and the selectivity of *rac*-**13e** were clearly reduced, compared to *rac*-**13d**.

Figure 3



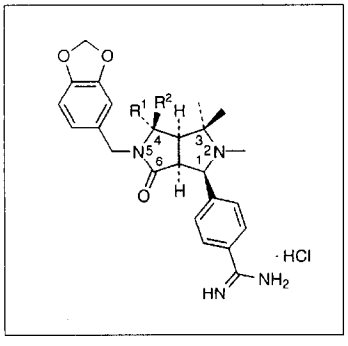
Synthesis of the inhibitors *rac*-**20b–f**. (i) Li[Et_3BH], THF, −78→0°C, 30 min. (ii) Toluene-4-sulfinic acid, $CaCl_2$, CH_2Cl_2, room temperature (r.t.), 24 h, 73%. (iii) $ZnCl_2$, RMgCl or RMgBr, CH_2Cl_2, 0°C→r.t., 13–36 h, 48% (*rac*-**18c**), 52% (*rac*-**18d**), 21% (*rac*-**18e**), 90% (*rac*-**18f**). (iv) CuCN, dimethylformamide (DMF), reflux, 17–30 h, 31% (*rac*-**19b**, yield from *rac*-**17**), 71% (*rac*-**19c**), 56% (*rac*-**19d**), 66% (*rac*-**19e**), 61% (*rac*-**19f**). (v) CH_3OH, HCl (g), $CHCl_3$, 4°C, 24 h. (vi) NH_3, CH_3OH, 65°C, 3.5 h, 78% (*rac*-**20b**), 62% (*rac*-**20c**), 79% (*rac*-**20d**), 72% (*rac*-**20e**), 89% (*rac*-**20f**).

Inhibitors lacking the 'lower' C=O group at C(6) were, as expected, less active than the imide *rac*-**21** (Table 2), because they were unable to form a hydrogen bond to the backbone NH of Gly216 (see Fig. 1). This hydrogen bond is not essential for an inhibitor to be active, because the binding affinities of imide *rac*-**21** and lactam *rac*-**9** differ by only $\Delta\Delta G_{21\rightarrow 9} = +0.8\,kcal\,mol^{-1}$. This value is in good agreement with data from Fersht *et al.* [16] who observed losses in the free energy of binding of 0.5–1.5 kcal mol^{-1} by deleting uncharged hydrogen bonds in enzyme–substrate complexes using site-directed mutagenesis. Lactam *rac*-**16**, with an *endo*-methyl group at C(6), still showed some affinity towards thrombin and trypsin; we cannot rule out the possibility that this inhibitor has a different binding mode, however.

The inhibitors with a tricyclic core structure (Table 3) are much more active and selective overall than those with a bicyclic one (Table 2). The three lactams with an *exo*-ethyl (*rac*-**20b**), an *exo*-cyclopropyl group (*rac*-**20c**) and an *exo*-isopropyl group (*rac*-**20d**) at C(1) showed high and similar binding affinities for thrombin. As in the bicyclic series, the inhibitor with the *exo*-isopropyl substituent (*rac*-**20d**) was the most selective. This compound ($K_i = 13$ nM) was

Table 1

Binding affinities of thrombin inhibitors with bicyclic core structures modified at C(4) and their selectivity with respect to trypsin.

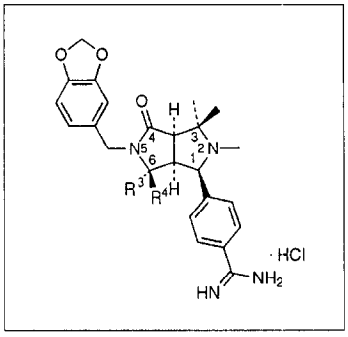


Compound	R^3	R^4	K_i (μM) Thrombin	K_i (μM) Trypsin	Selectivity*	$\Delta\Delta G^{\dagger}$ (kcal mol^{-1})
rac-**21**	O		0.5	1.1	2.2	−1.5
rac-**8**	H	H	6.2	31	5	0
rac-**13a**	Methyl	H	2.0	64	32	−0.7
rac-**13b**	Ethyl	H	0.095	7.3	77	−2.5
rac-**13d**	Isopropyl	H	0.030	6.7	223	−3.2
rac-**13e**	Cyclohexyl	H	0.86	33	38	−1.2

*K_i (trypsin) / K_i (thrombin). †Contributions of substituents at C(4) to the binding free energy at 298 K. $\Delta\Delta G = \Delta G$(*rac*-**8**)$-\Delta G(x)$; x: racemic compounds **21 and 13a,b,d,e**. $\Delta G = -RT \ln K_i$.

Table 2

Binding affinities of thrombin inhibitors with bicyclic core structure modified at C(6) and their selectivity with respect to trypsin.



Compound	R^3	R^4	K_i (μM) Thrombin	K_i (μM) Trypsin	Selectivity*	ΔΔG† (kcal mol^{-1})
rac-**21**	O		0.5	1.1	2.2	0
rac-**9**	H	H	2.0	64	32	+0.8
rac-**16**	H	Methyl	9.2	40	4.3	+1.7

*K_i (trypsin) / K_i (thrombin). †Destabilization ΔΔG of the thrombin–inhibitor complex by removal or replacement of one carbonyl group of *rac*-**21**. ΔΔG = ΔG(*rac*-**21**) − ΔG(x); x: racemic compounds **9** and **16**.

not only selective for thrombin over trypsin [K_i(trypsin) = 9.9 μM], but also showed high selectivity when tested for binding to the coagulation factors Xa (K_i = 140 μM) and VIIa (K_i > 500 μM).

Inhibitor *rac*-**20d** is a mimetic of the natural thrombin substrate fibrinogen. This protein binds in the D-pocket with the phenyl group of a phenylalanine (*rac*-**20d** uses a piperonyl group), in the P-pocket with the isopropyl group of a valine (*rac*-**20d** also uses an isopropyl group), and in the S1-pocket with the guanidinium sidechain of an arginine residue (*rac*-**20d** uses an amidinium group) [17]. A substantial loss in affinity was again detected for inhibitors with larger substituents at C(4) — such as *exo*-cyclohexyl (*rac*-**20e**) and *exo*-phenyl (*rac*-**20f**). These groups are too voluminous to be incorporated into the P-pocket of thrombin.

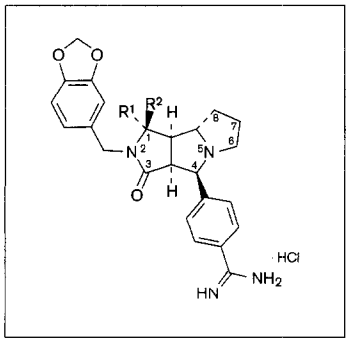
The observation that small hydrophobic moieties are preferentially bound in the P-pocket of thrombin confirms the results from the literature [15]. Obviously, the loop in thrombin that forms the P-pocket is rather rigid and can only adapt to larger groups in the inhibitor to a limited extent. A methyl group, however, is too small to fill the space available in this pocket.

Optical resolution of *rac*-20d

All biological activities discussed above refer to racemic mixtures of the inhibitors. Based on the molecular modeling experiments, we expected a very large difference in the affinity of the two enantiomers towards thrombin. Consequently, we carried out the optical resolution (Fig. 4) of the most selective inhibitor *rac*-**20d**. Imide *rac*-**3** was reduced to the hydroxylactam and subsequently treated with methyl D-(−)-mandelate and a catalytic amount of pyridinium-4-toluenesulfonate in refluxing toluene. The resulting diastereoisomeric ethers **23** and **24** were separated by column chromatography over silica gel and converted into the enantiomerically pure sulfones (+)-**17** (enantiomeric purity 96.5%) and (−)-**17** (enantiomeric

Table 3

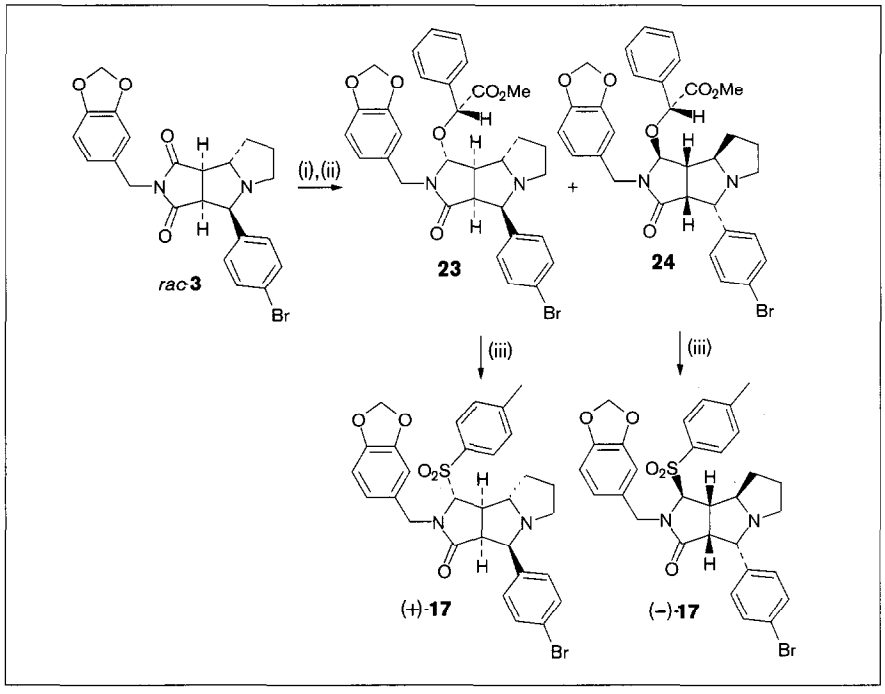
Binding affinities of thrombin inhibitors with tricyclic core structure modified at C(1) and their selectivity with respect to trypsin.



Compound	R^1	R^2	K_i (μM) Thrombin	K_i (μM) Trypsin	Selectivity*	ΔΔG† (kcal mol^{-1})
rac-**1**‡	O		0.09	0.70	7.8	0
rac-**20b**	Ethyl	H	0.0081	1.7	210	−1.4
rac-**20c**	Cyclopropyl	H	0.010	2.3	230	−1.3
rac-**20d**	Isopropyl	H	0.013	9.9	760	−1.1
rac-**20e**	Cyclohexyl	H	1.7	4.4	2.6	+1.7
rac-**20f**	Phenyl	H	1.4	5.2	3.6	+1.6

*K_i (trypsin) / K_i (thrombin). †Stabilization or destabilization ΔΔG of the thrombin–inhibitor complex by replacement of one carbonyl group of **1**. ΔΔG = ΔG(*rac*-**1**) − ΔG(x); x: racemic compounds **20b–f**. ‡from [9].

Figure 4



Optical resolution of sulfone *rac*-**17**. (i) $Li[Et_3BH]$, THF, −78→0°C, 30 min. (ii) Methyl D-(−)-mandelate, pyridinium-4-toluenesulfonate, toluene, reflux, 6 h, 61% (**23**), 64% (**24**). (iii) Toluene-4-sulfinic acid, $CaCl_2$, CH_2Cl_2, r.t., 12 h. (+)-**17**: 73%. (−)-**17**: 91%.

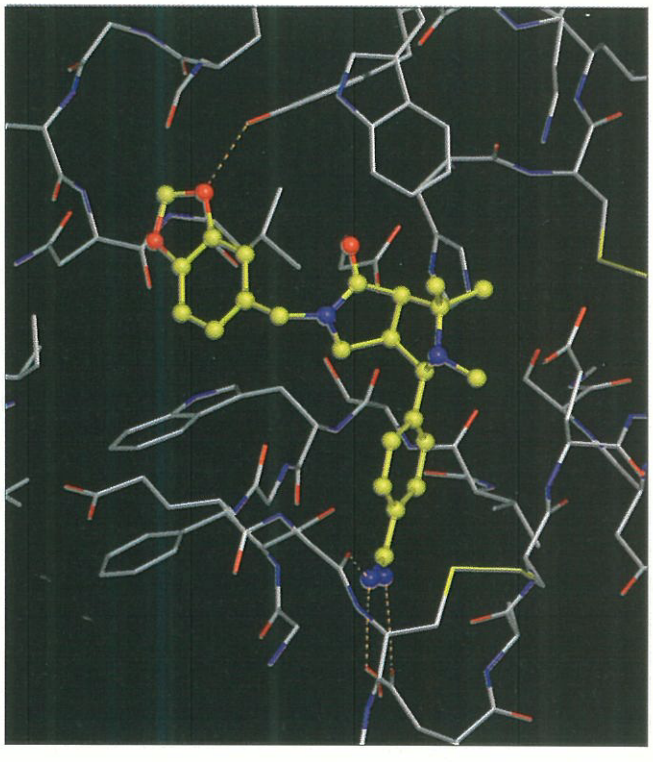
purity >99.5%). A further purification of (+)-**17** was achieved by recrystallization from ethyl acetate/hexane. The very insoluble racemate crystallized out, leaving the remaining soluble (+)-**17** with an enantiomeric purity of higher than 99:1.

The enantiomeric sulfones (+)-**17** and (−)-**17** were then converted into the corresponding amidinium salts (+)-**20d** and (−)-**20d** by the procedure used for the preparation of the racemic mixture. Enantiomer (+)-**20d** bound to thrombin with $K_i = 7$ nM and a selectivity K_i(trypsin)/K_i(thrombin) of 740. It was therefore 800 times more active than (−)-**20d** ($K_i = 5.6\,\mu$M, selectivity 21). The more potent enantiomer was the one found in the crystal structure of the complex formed between thrombin and (±)-**20d** (see below), and, consequently, we assigned the (1*R*,3a*S*,4*R*, 8a*S*,8b*R*)-configuration to (+)-**20d** (numbering according to Fig. 3).

Confirmation of the binding modes by X-ray crystallography

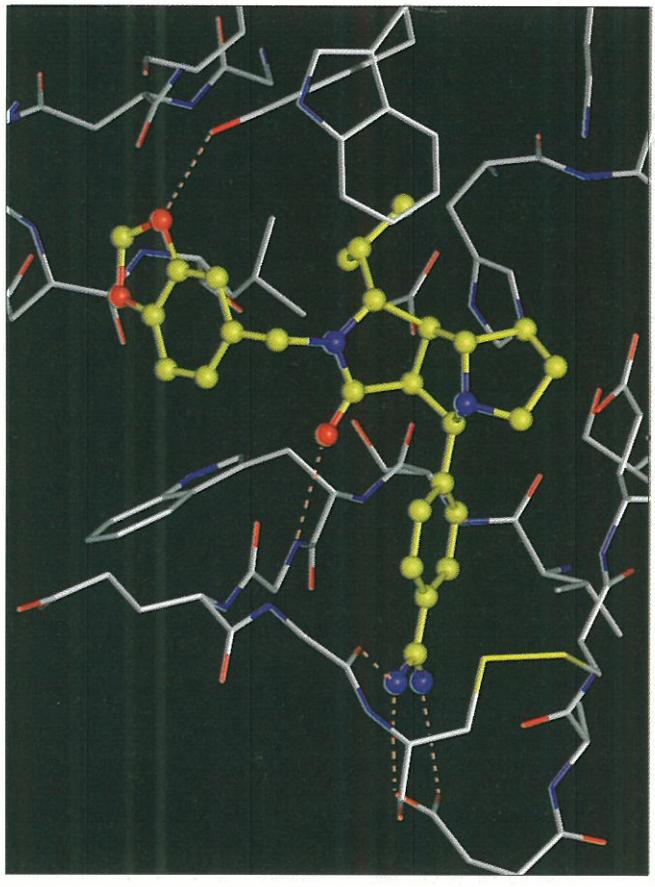
The X-ray crystal structures of the complexes of lactams *ent*-**9** (resolution 2.1 Å) and (+)-**20d** (resolution 1.93 Å) with thrombin (Figs 5,6) showed that the more potent inhibitor enantiomers were incorporated into the complex and adopted the same binding mode in the enzyme active site as did the imide *ent*-**1** [9]. This was as expected from molecular modeling studies. The observation of a nearly identical binding mode in three crystal structures, as a result of the high degree of structural preorganization of the inhibitors, fully validates the correlations between

Figure 5



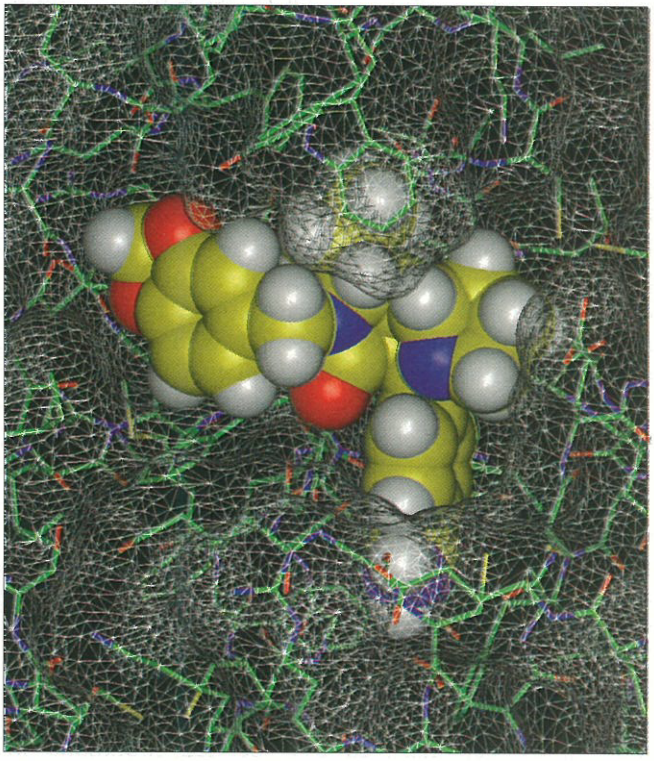
The active site region in the X-ray crystal structure of the complex between thrombin and *ent*-**9**. Hydrogen bonds are represented as dashed lines.

Figure 6



The active site region in the X-ray crystal structure of the complex between thrombin and (+)-**20d**. Hydrogen bonds are represented as dashed lines.

Figure 7



The active site region in the X-ray crystal structure of the complex between thrombin and (+)-**20d**. Space filling representation of (+)-**20d** and solvent-accessible surface of thrombin.

changes in binding free energy and the bonding differences between individual substituents that were drawn in the molecular recognition analysis above.

Interestingly, removal of the C=O group at C(6) upon passing from imide *ent*-**1** (Fig. 1) to lactam *ent*-**9** (Fig. 5) did not induce a significant positional shift of the bicyclic core structure within the protein, although the hydrogen bond to Gly216 was lacking. Apparently, the backbone NH of Gly216 does not require strong solvation through hydrogen bonding to the inhibitor, as was also evident from the modest difference in the free energy of binding between the complexes of imide *rac*-**21** and lactam *rac*-**9** (Table 2). The low energetic costs for the loss of this intermolecular hydrogen bond can be explained by the electrostatic stabilization of the partially positively-charged hydrogen atom in the NH group by the π-electrons of the phenylamidinium residue of *ent*-**9** which is oriented cofacially to the planar –CO–C–NH– moiety of Gly216 at a distance of ~3.6 Å. In the planar glycine residue, dipolar compensation between the antiparallel C=O and N–H dipoles should provide substantial internal solvation. The advantages of the hydrogen bond between the lower carbonyl oxygen atom at C(6) of *ent*-**1** and (+)-**20d** and the NH of Gly216 are partially compensated for by repulsive secondary electrostatic interactions [18] between this oxygen atom and the carbonyl oxygen atom of Gly216, located in close contact at 2.9 Å. Also, favorable dispersion interactions between the $H_2C(6)$ group in *ent*-**9** and the CH–CH_2 moiety of Trp215, as well as the corresponding hydrophobic desolvation, should partially compensate for the loss of the hydrogen bond to Gly216.

The isopropyl substituent of (+)-**20d** fits well into the P-pocket as shown in Figures 6 and 7. Its pro-*S* methyl group (on the left side) has close C···C contacts (3.3–4.1 Å) with an aromatic carbon atom of Tyr60A, the CH_2 group of His57, and a CH_3 group of Leu99. Furthermore, the pro-*S*-CH_3–CH moiety packs intramolecularly against the piperonyl group. Hydrophobic packing of inhibitor substituents that bind into the P-pocket and D-pocket of the enzyme is often seen in thrombin–inhibitor complexes [5]. The pro-*R* methyl group has short C···C contacts with aromatic carbon atoms of Tyr60A and Trp60D, that form the 'ceiling' and the right wall of the P-pocket, as well as with a heterocyclic carbon atom of His57.

Significance

Thrombin is a trypsin-like serine protease that is central in the processes of hemostasis and thrombosis. The therapeutic control of thrombin activity by direct acting, selective inhibitors could have advantages over the currently used antithrombotics heparin and coumarins. In this work, we present the structure-based design and synthesis of a series of potent and selective novel thrombin inhibitors. These compounds have a rigid bicyclic or tricyclic core structure and are highly preorganized for binding to thrombin. The nonpeptidic nature of this class of inhibitors makes them interesting lead molecules for the continuing search for compounds with optimal pharmacological properties. Future analogs have the potential to be useful for the treatment and prevention of thrombotic diseases. The identical binding mode of the inhibitors at the thrombin active site, which is a direct consequence of their high degree of conformational preorganization by the bicyclic and tricyclic core structures, provided the basis for sound biological molecular recognition studies. Thus, measured changes in the free energy of binding could be directly correlated with the bonding contributions of individual substituents of the inhibitors and valuable information on the magnitude of individual intermolecular interactions was obtained. This study demonstrates that defined mutations in highly preorganized inhibitors provide an attractive alternative to site-directed mutagenesis in exploring molecular recognition phenomena at enzyme active sites.

Materials and methods

Molecular modeling

The design of the target molecules was carried out on a Silicon Graphics Crimson workstation using the molecular modeling program MOLOC (F. Hoffmann-La Roche) [10]. Modeling procedure: the proposed inhibitor was minimized separately and docked manually into its expected binding site. The coordinates of thrombin were fixed and the inhibitor was minimized inside the enzyme.

Analytical characterization of rac-***20d***

Melting point (m.p.) 210–215°C. ^{1}H-NMR (200 MHz, $(CD_3)_2SO$, see Fig. 3 for the numbering of atoms): δ 0.65 (d, J = 6.4, 3H, CH_3); 0.89 (d, J = 6.7, 3H, CH_3); 1.63 (m, 2H, H–C(7) and H–C(8)); 1.90 (m, 2H, H–C(7) and H–C(8)); 2.09 (m, 1H, CH–C(1)); 2.37–2.62 (m, 2H, H–C(8b) and H–C(6)); 2.82 (m, 1H, H–C(6)); 3.17 (m, 2H, H–C(1) and H–C(8a)); 3.36 (m, 1H, H–C(3a)); 3.77, 4.53 (AB, J = 14.9, 2H, CH_2–N(2)); 4.18 (d, J = 7.5, 1H, H–C(4)); 6.02 (m, 2H, OCH_2O); 6.70, 6.89 (AB, J=7.9, 2H, piperonyl aromatic); 6.74 (s, 1H, piperonyl aromatic); 7.56, 7.76 (AB, J=8.3, 4H, 4-amidiniumphenyl); 9.13 (s, 2H, $C(NHH)_2$); 9.34 (s, 2H, $C(NHH)_2$). Desorption electron impact high resolution mass spectrum: calc'd for $C_{27}H_{32}N_4O_3$ ([M–HCl]$^+$ 460.2474; found 460.2488.

Analytical characterization of optically pure compounds

The enantiomeric purity of the sulfones (+)-**17** and (–)-**17** was determined by analytical HPLC (AcOEt : hexane 1:1) on a commercial (*S,S*)-Whelk-O1 column (250 mm × 4.6 mm I.D.) from Regis, 8210 Austin Ave, Morton Grove, IL 60053, USA [19]; detection at 254 nm; separation factor α=1.8. (+)-**17**: $[\alpha]_D^{25}$+243 (c 1.00, $CHCl_3$), enantiomeric purity >99:1.(–)-**17**: $[a]_D^{25}$–237 (c 1.00, $CHCl_3$), enantiomeric purity >99.5:0.5. (+)-**20d**: decomposition >150°C, $[a]_D^{25}$+142 (c 1.00, $CHCl_3$); (–)-**20d**: decomposition >150°C, $[a]_D^{25}$–145 (c 0.99, $CHCl_3$).

Full experimental protocols for the syntheses of the inhibitors described in this work will be provided in a full paper. All new compounds reported here were fully characterized by mass spectrometry, ^{13}C-NMR, ^{1}H-NMR, IR, and elemental analysis or high resolution mass spectra.

X-ray crystal structure determination

Crystals of C2 symmetry of human α-thrombin were grown in the presence of a carboxy-terminal hirudin peptide [des-amino Asp55 hirudin 55–65] as inhibitor in the anion-binding exosite, but with no active-site inhibitor. Compounds *rac*-**9** and *rac*-**20d** were diffused into crystals overnight at a concentration of 1 mM. Diffraction data were measured using a 30 cm image plate (Marresearch) on an FR591 X-ray generator (Delft Instruments) equipped with double focusing mirrors (Supper) run at 3.5 kW. Exposure times were 300 s for 0.5° frames and complete data were obtained in a 120° scan. Only the higher-affinity enantiomer of the two inhibitors was found in the X-ray crystal structure. For the complex of *ent*-**9**, the detector distance was 160 mm, the limiting resolution 2.1 Å, and the unit cell dimensions a=71.99 Å, b=72.52 Å, c=73.20 Å, and β=100.80°. The merging R-factor on intensities was 3.4% and I/ΣI=3.1 in the outermost data shell. Refinement with X-plor gave final crystallographic R-factors of 14.9% (working) and 19.8% (free), with rms (root mean square) bond errors of 0.01 Å and rms angle errors of 1.86°. For the complex of (+)-**20d** the detector distance was 140 mm, the limiting resolution 1.93 Å, and the unit cell dimensions were a=71.95 Å, b=72.41 Å, c=73.41 Å, and β=101.037°. Refinement with X-plor gave a final crystallographic R-factor of 15.9%, with rms bond errors of 0.01 Å and rms angle errors of 1.84°.

Determination of inhibition constants

The affinity of the thrombin inhibitors was determined according to [14,15] (chromogenic substrate S-2238). An exhaustive protocol of the binding assay identical to the one used in this study is provided in [15].

Acknowledgements

We thank the Fonds der Deutschen Chemischen Industrie for a Kekulé-Mobilitätsstipendium (U.O.) and F. Hoffmann-La Roche for supporting this study.

References

1. Scully, M.F. (1992). The biochemistry of blood clotting: the digestion of a liquid to form a solid. In *Essays in Biochemistry*. (Tipton, K.F., ed.), pp. 17–36, Portland Press, London, UK.
2. Lyle, T.A. (1993). Small-molecule inhibitors of thrombin. *Perspect. Drug Disc. Des.* **1**, 453–460.
3. Balasubramanian, B.N. (ed.) (1995). Advances in the design and development of thrombin inhibitors. *Bioorg. Med. Chem. Symp. 5* **3**, 999–1156.
4. Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S.R. & Hofsteenge, J. (1989). The refined 1.9 Å crystal structure of human α-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J.* **8**, 3467–3475.
5. Banner, D.W. & Hadvary, P. (1991). Crystallographic analysis at 3.0-Å resolution of the binding to human thrombin of four active site-directed inhibitors. *J. Biol. Chem.* **266**, 20085–20093.
6. Reich, S.H. & Webber, S.E. (1993). Structure-based drug design (SBDD): every structure tells a story... *Perspect. Drug Disc. Des.* **1**, 371–390.
7. Lam, P.Y.S., *et al.*, & Erickson-Viitanen, S. (1994). Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* **263**, 380–384.
8. Greer, J., Erickson, J.W., Baldwin, J.J. & Varney, M.D. (1994). Application of the three-dimensional structures of protein target molecules in structure-based drug design. *J. Med. Chem.* **37**, 1035–1054.
9. Obst, U., Gramlich, V., Diederich, F., Weber, L. & Banner, D.W. (1995). Design of novel, nonpeptidic thrombin inhibitors and structure of a thrombin-inhibitor complex. *Angew. Chem. Int. Ed. Engl.* **34**, 1739–1742.
10. Gerber, P.R. & Müller, K. (1995). MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput. Aided Mol. Design* **9**, 251–268.

11. Grigg, R., Idle, J., McMeekin, P. & Vipond, D. (1987). The decarboxylative route to azomethine ylides. Mechanism of 1,3-dipole formation. *J. Chem. Soc. Commun.* 49–51.
12. Richter, P., Kazmirowski, H.-G., Wagner, G. & Garbe, C. (1973). Synthesen von 4-Amidinophenylbrenztraubensäure (APPA) (Translation: Synthesis of 4-amidino phenyl pyruvic acid.) *Pharmazie* **28**, 585–591.
13. Brown, D.S., Charreau, P., Hansson, T. & Ley, S.V. (1991). Substitution reactions of 2-phenylsulphonyl-piperidines and -pyrrolidines with carbon nucleophiles: synthesis of the pyrrolidine alkaloids norruspoline and ruspolinone. *Tetrahedron* **47**, 1311–1328.
14. Lottenberg, R., Hall, J.A., Fenton, J.W. & Jackson, C.M. (1982). The action of thrombin on peptide *p*-nitroanilide substrates: hydrolysis of Tos-Gly-Pro-Arg-pNA and D-Phe-Pip-Arg-pNA by human α- and γ- and bovine α- and β-thrombins. *Thromb. Res.* **28**, 313–332.
15. Hilpert, K., *et al.*, & Van de Waterbeemd, H. (1994). Design and synthesis of potent and highly selective thrombin inhibitors. *J. Med. Chem.* **37**, 3889–3901.
16. Fersht, A.R., *et al.*, & Winter, G. (1985). Hydrogen bonding and biological specificity analysed by protein engineering. *Nature* **314**, 235–238.
17. Martin, P.D., Robertson, W., Turk, D., Huber, R., Bode, W. & Edwards, B.F.P. (1992). The structure of residues 7–16 of the Aα-chain of human fibrinogen bound to bovine thrombin at 2.3-Å resolution. *J. Biol. Chem.* **267**, 7911–7920.
18. Jorgensen, W.L. & Severance, D.L. (1991). Chemical chameleons: hydrogen bonding with imides and lactams in chloroform. *J. Am. Chem. Soc.* **113**, 209–216.
19. Pirkle, W.H., Welch, C.J. & Lamm, D. (1992). Design, synthesis, and evaluation of an improved enantioselective naproxen selector. *J. Org. Chem.* **57**, 3854–3860.