doi:10.1016/j.jmb.2009.04.051 J. Mol. Biol. (2009) 390, 56–69

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

More than a Simple Lipophilic Contact: A Detailed Thermodynamic Analysis of Nonbasic Residues in the S1 Pocket of Thrombin

Bernhard Baum[1], Menshawy Mohamed[2], Mohamed Zayed[2], Christof Gerlach[1], Andreas Heine[1], David Hangauer[2]* and Gerhard Klebe[1]*

[1]*Department of Pharmaceutical Chemistry, Philipps-University Marburg, Marbacher Weg 6, 35032 Marburg, Germany*

[2]*Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, NY 14260, USA*

Received 16 January 2009; received in revised form 23 April 2009; accepted 23 April 2009
Available online 3 May 2009

The field of medicinal chemistry aims to design and optimize small molecule leads into drug candidates that may positively interfere with pathological disease situations in humans or combat the growth of infective pathogens. From the plethora of crystal structures of protein–inhibitor complexes we have learned how molecules recognize each other geometrically, but we still have rather superficial understanding of why they bind to each other. This contribution surveys a series of 26 thrombin inhibitors with small systematic structural differences to elucidate the rationale for their widely deviating binding affinity from 185 μM to 4 nM as recorded by enzyme kinetic measurements. Five well-resolved (resolution 2.30 – 1.47 Å) crystal structures of thrombin–inhibitor complexes and an *apo*-structure of the uncomplexed enzyme (1.50 Å) are correlated with thermodynamic data recorded by isothermal titration calorimetry with 12 selected inhibitors from the series. Taking solubility data into account, the variation in physicochemical properties allows conclusions to be reached about the relative importance of the enthalpic binding features as well as to estimate the importance of the parameters more difficult to capture, such as residual ligand entropy and desolvation properties. The collected data reveal a comprehensive picture of the thermodynamic signature that explains the so far poorly understood attractive force experienced by *m*-chloro-benzylamides to thrombin.

Keywords: structure-based drug design; thrombin; thermodynamic driving forces; residual ligand entropy; chlorine–π interaction

Edited by J. E. Ladbury

Introduction

Nearly two decades of drug research on small molecules inhibiting the serine protease thrombin, an established target for the prevention of cardiovascular diseases, have revealed a plethora of highly potent drug candidates.[1] Most of the early development programs followed a strategy to address the S1 specificity pocket of the enzyme with basic molecular moieties to mimic the arginine side chain of the natural substrate fibrinopeptide A.[2] However, as a consequence of being positively charged under physiological conditions, they generally have poor pharmacokinetic properties. In order to address this deficiency, sophisticated prodrug strategies have been developed to achieve sufficient bioavailabilty.[3] Ongoing efforts to produce therapeutically useful orally available thrombin inhibitors resulted in the discovery of the uncharged *m*-chloro-benzylamides as promising alternatives to the prodrug strategy.[4] They are assumed to form a lipophilic contact in the S1 pocket,[5–8] which has the polar carboxylate group of Asp189 at its far end. Therefore the rationalization put forward[5–8] to explain the pronounced binding affinities of the *m*-chloro derivatives as predominantly hydrophobic binding appears unsatisfactory. Detailed investigations with respect to the binding thermodynamics of these extremely promising

**Corresponding authors.* E-mail addresses: hangauer@buffalo.edu; klebe@staff.uni-marburg.de.

Abbreviations used: ITC, isothermal titration calorimetry; DCM, dichloromethane.

ligands to elucidate their binding characteristics are described in the present contribution.

We present a detailed study, including X-ray crystallography and isothermal titration calorimetry, on a congeneric series of thrombin inhibitors to find possible explanations for these so far poorly understood structure–activity relationships.

An electrostatic attractive force resulting from a polarized chlorine atom above a tyrosine π-system has been demonstrated recently for chloro-aromatic inhibitors addressing the structural related S1 pocket of coagulation factor Xa.[9–13] An analysis of crystal structures from the Protein Data Bank along with *ab initio* calculations showed the Cl–π interaction as a clearly attractive force, which was assigned to the dispersion force as major source of attraction.[14] A better understanding of the electrostatic component of this interaction may also make medicinal chemists more aware of the possible nature of such interactions and the transfer of this knowledge to other related and ongoing drug design projects. The astonishingly high ligand efficiency of triclosane to its molecular target FabI,[15–18] a ligand that intuitively contradicts our keen sense for an optimal lead compound, re-emphasizes the necessity to better understand the driving forces of chloro-aromatic interactions in protein–ligand complex formation.

The characterization of the determinants for molecular recognition of small molecules by a macromolecular target is an essential requirement for any drug development project that intends to start from rational grounds.[19] Whereas the foundation of these mostly non-covalent interactions is reasonably well understood, it is still a major challenge to obtain detailed knowledge about the individual thermodynamic parameters and their respective contributions to binding affinity. Nevertheless, studies providing this detailed information would allow us to predict more precisely ligand binding affinity as a prerequisite for successful *in silico* screening.[20–22]

The interplay of intra- and intermolecular forces between ligand, receptor and solute makes the binding event a process of challenging thermodynamic complexity. A promising approach to unravelling this complexity is to perform systematic studies on interacting molecules with only minor differences.[23] Few systematic studies are available,[24–26] and they often demonstrate paradoxical or unexpected results.[27] They also emphasize that entropy can significantly affect the Gibbs free energy of binding, and that entropy is notoriously difficult to predict.[28] If only the overall binding affinity, as determined by a typical bio-assay, is taken into account, the danger exists that a superficial and oversimplified

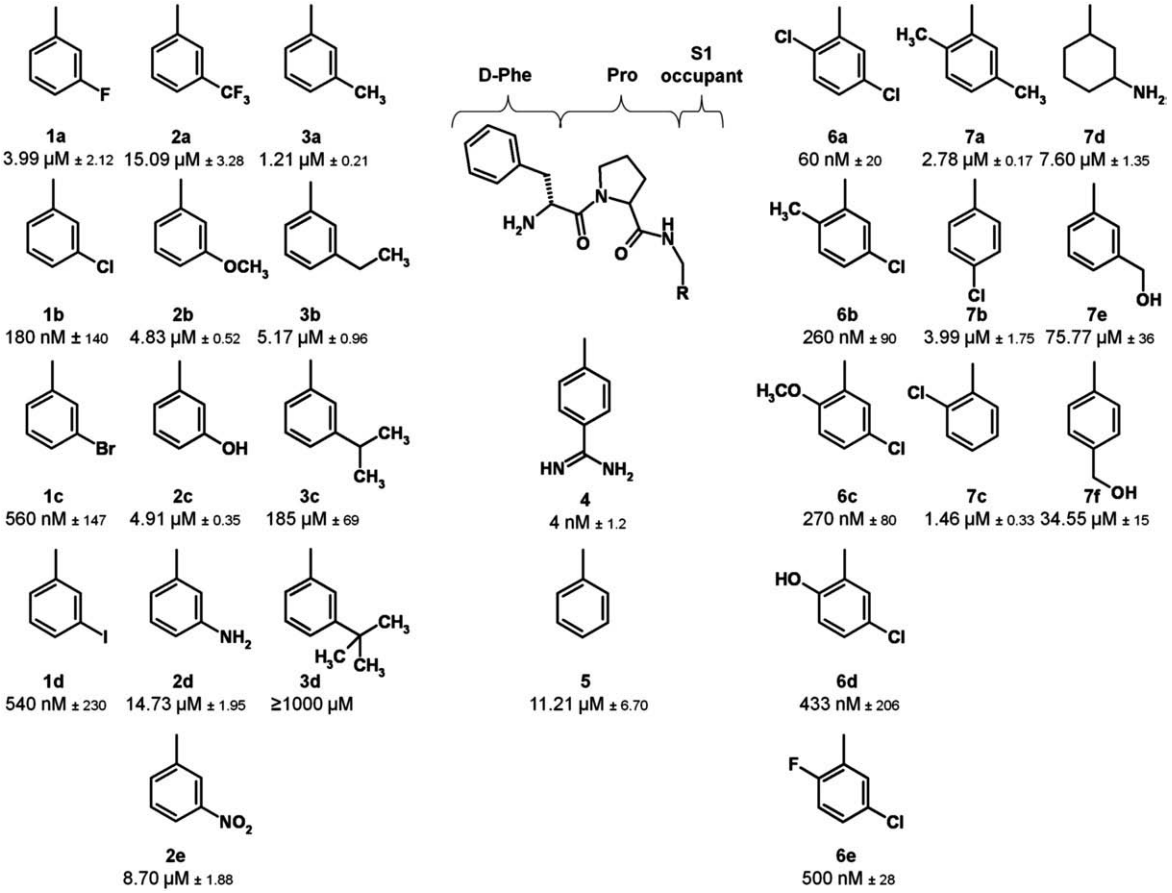

Fig. 1. Chemical formulae of the inhibitors discussed here. Inhibitory potency towards human thrombin is given as kinetic inhibition constant K_i as the mean and standard deviation of at least three measurements. Curly brackets indicate the three parts of each molecule for which *B*-factors from crystal structures were evaluated separately (Table 1).

interpretation of binding energetics will be made.[29] Consequently, multiple biophysical methods should be utilized in order to clarify the energetics of the ligand binding process in greater depth.

The serine protease thrombin[30] is an ideal model system for performing fundamental studies aimed at providing a better understanding of the energetics of the ligand–protein binding process because its active site is composed of a set of distinct, well-defined, binding pockets that can easily be addressed.[31] Crystallography can be performed routinely and the large quantities of enzyme necessary for isothermal titration calorimetry (ITC) are available. Furthermore, thrombin is a target familiar to many researchers in the enzyme inhibitor field and, therefore, using it as a model system facilitates an appreciation of the information obtained as well as its application to other enzyme systems.

Here, we present a "thermodynamic mapping" of a series of D-Phe-Pro-based thrombin inhibitors (Fig. 1), that differ only in the portion binding to the S1 pocket. This ligand series allows us to follow changes in enthalpic and entropic contributions to the Gibbs free energy of binding in a stepwise fashion. A series of 26 ligands has been synthesized and their binding affinities towards human thrombin were first determined in a kinetic assay. For 12 of the most interesting compounds, the Gibbs free energy of binding ΔG^0 could be factorized by ITC into an enthalpic ΔH^0 and entropic contribution $-T\Delta S^0$. As a reference state for thrombin before the binding event, a crystal structure of the uncomplexed enzyme was determined, showing an extended hydration pattern of the active site. Five inhibitors with wide-ranging binding affinity, even though they are structurally very similar, have been investigated by X-ray crystallography. Their binding modes elucidate the interaction pattern and hydration in the S1 specificity pocket and allow conclusions about residual ligand mobility in light of their relative differences in the temperature factor.

When combined with solubility and desolvation energy considerations, the data provide a more complete analysis of the driving forces for inhibitor binding, in which unexpected electrostatic attraction and residual entropy play important roles.

Results and Discussion

Analysis of crystal structures

X-ray crystallography of uncomplexed thrombin resulted in a structure with the active site of the enzyme in its hydrated state. Eight structural water molecules could be assigned to peaks in the difference electron density for the region that is usually occupied by the inhibitor series under consideration. The assignment of these water molecules during refinement was made on the basis of the initial difference electron density and on expedient temperature factors. Although the complete solution hydration pattern is not fully represented by the crystallographically observed hydration water molecules, the eight observed water molecules within the inhibitor binding cavity provide a reasonable approximation of the putative hydrogen bonding network that has to be disrupted upon inhibitor binding.

Furthermore, the *apo*-structure provides insights into the effects of inhibitor binding on the conformation of the amino acids of the active site. The major structurally observable difference between the uncomplexed enzyme and the inhibitor-bound state is that the side chain of Leu99 at the side wall of the S3 pocket could be refined in a double conformation when none of the ligands in the present study was

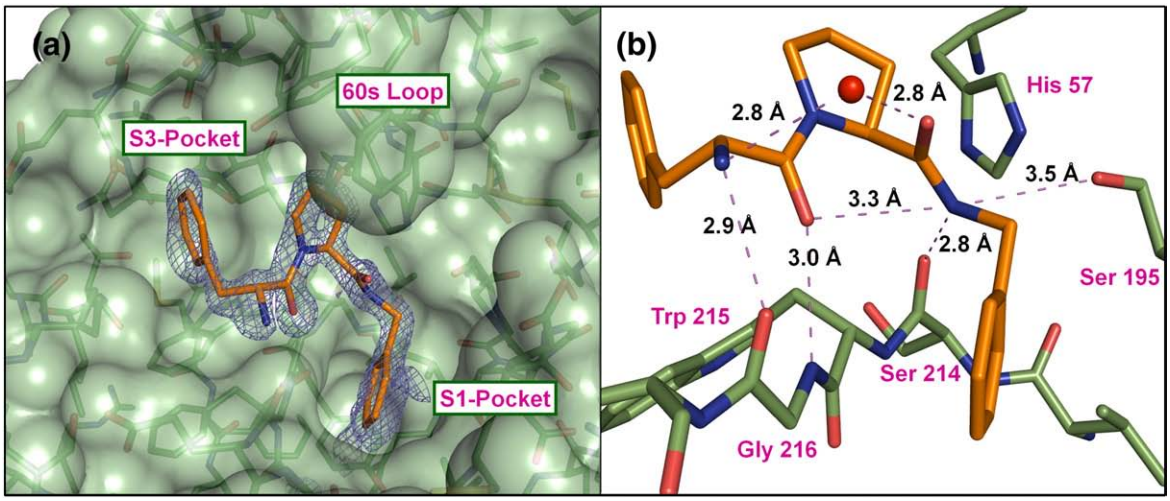

Fig. 2. Inhibitor 5 in complex with human thrombin inhibited at the exosite by hirugen. (a) View of the active site with the solvent-accessible surface of the thrombin heavy chain in pale green and the inhibitor in orange, heteroatoms are colour-coded. The $F_o - F_c$ difference electron density for 5 is depicted in blue at 1.5σ. (b) Active site residues are represented by green sticks, ligand 5 is shown in orange. The red sphere represents the water molecule involved in a hydrogen bonding network. Hydrogen bonds are indicated as broken magenta lines.

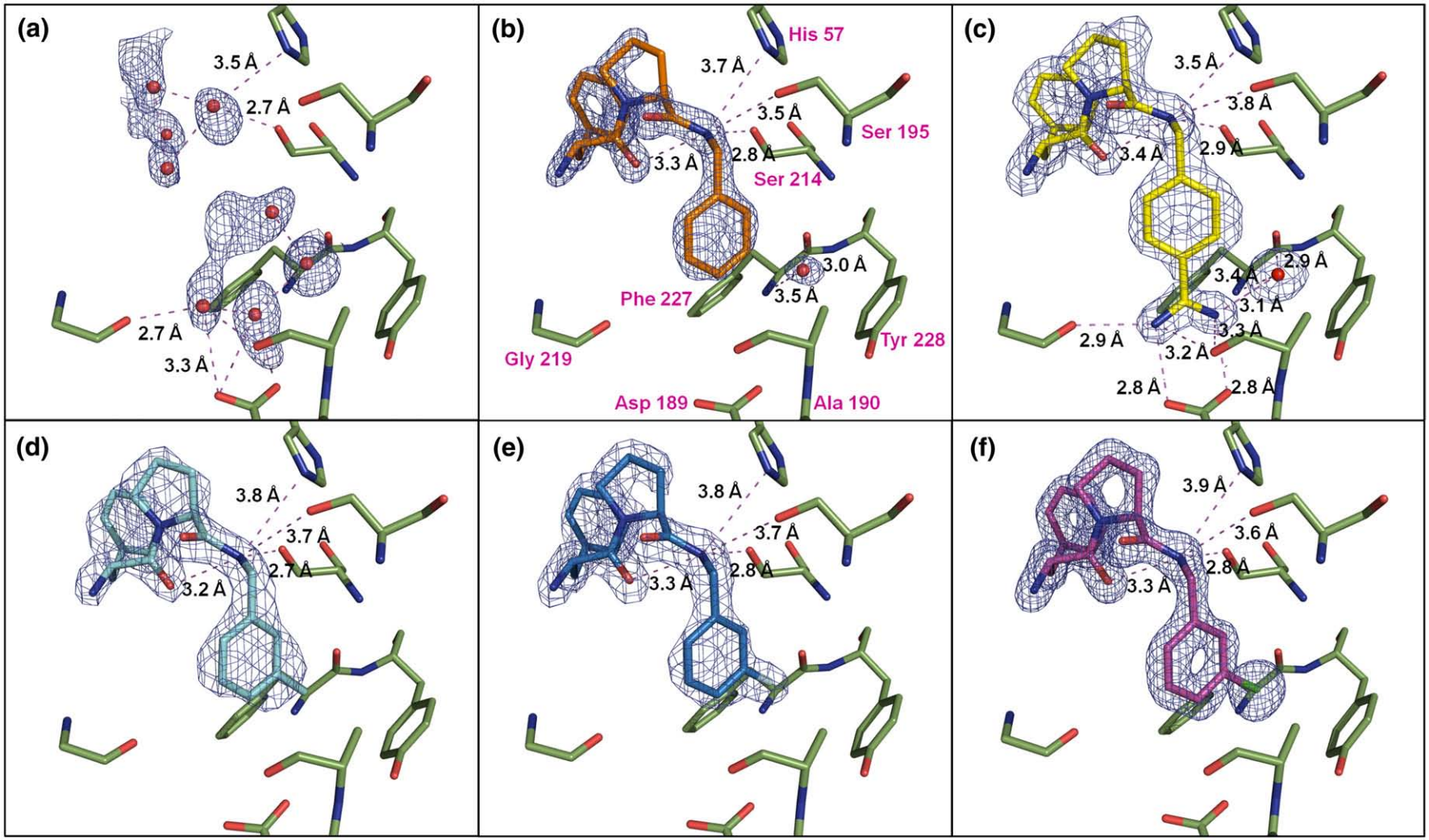

Fig. 3. View of the S1 specificity pocket of the uncomplexed enzyme (a) and with bound inhibitors **5** (b), **4** (c), **3**a (d), **1**a (e) and **1**b (f). The $F_o - F_c$ difference electron density for all ligands and water in the pocket is depicted in blue at 2σ. Important distances to residues of the binding pocket are indicated as broken magenta lines.

occupying this pocket. This indicates a higher degree of thermal motion for this amino acid, relative to the others, in the uncomplexed hydrophobic pocket, which becomes restricted to one conformation upon inhibitor binding. As the loss of rotational degrees of freedom for Leu99 in the S3 pocket is the same for all inhibitors discussed here, its effect on the relative differences of the thermodynamics of binding can be neglected in the direct comparison of the inhibitors.

The crystal structures of inhibitors **1a**, **1b**, **3a**, **4** and **5** in complex with human thrombin show virtually the same binding mode for the D-Phe-Pro portion of the ligands (Fig. 2a). The phenyl ring occupies the hydrophobic S3 pocket, whereas the proline moiety fits perfectly well beneath Tyr60A and Trp60D of the 60 loop, as was analogously observed for a valine side chain in the natural substrate.[32] The central region of the ligands is involved in an elaborate hydrogen bonding network with the backbone of the enzyme, along with a captured structural water molecule (Fig. 2b). All of the ligands form a β-sheet like interaction with the backbone residues 214 – 216 of the protein, as was observed for the natural substrate.[32] The terminal primary amine and the adjacent carbonyl form hydrogen bonds with distances of 2.7 – 3.1 Å to the corresponding functionalities of Gly216. Due to the highly conserved interaction pattern for the common D-Phe-Pro portion of the ligands in this study, any difference in their binding energetics can be confidently attributed to differences in their respective side chains binding in the S1 pocket.

The interaction pattern experienced inside the S1 pocket differs among the inhibitors investigated (Fig. 3). The benzamidine **4** binds tightest with a K_i of 4 nM. It forms an almost symmetric salt-bridge to Asp189 at the bottom of the pocket. Simultaneously, a water molecule is locked in the pocket in exactly the same position as was found in the uncomplexed state, and mediates hydrogen bonding from the ligand amidino group to the backbone of Phe227 (Fig. 3c). This water molecule is displaced by the *meta*-substituents CH_3, F and Cl of the benzylamides **3a**, **1b** and **1a** (Fig. 3d–f). The orientation of the terminal phenyl moiety with respect to Tyr228 results in a rotation of 33.4° compared to the phenyl ring orientation of **4.** The unsubstituted inhibitor **5**, even though it lacks the *meta*-substituent, adopts the same ring orientation (Fig. 3b). Furthermore, a water molecule returns into the complex at the same position as in the uncomplexed S1 binding pocket and in the benzamidine complex (Fig. 3a and c), albeit the water molecule exhibits in this case a much higher temperature factor accompanying inhibitor **5**. The location within the S1 pocket seems optimal for accommodating a water molecule, because it is able to form multiple hydrogen bonds with the protein at this site. Interestingly, the chlorine atom of **1b** falls close to this position with short distances (3.4 Å) to the carbonyl groups of Phe227 and Trp215 (Figs. 3f and 4). Furthermore, the centre of the chlorine atom is located 4.5 Å above the centroid of Tyr228. With this centroid it forms an angle of 67.4° with respect to the plane through the phenolic ring. This geometry resembles the preferred

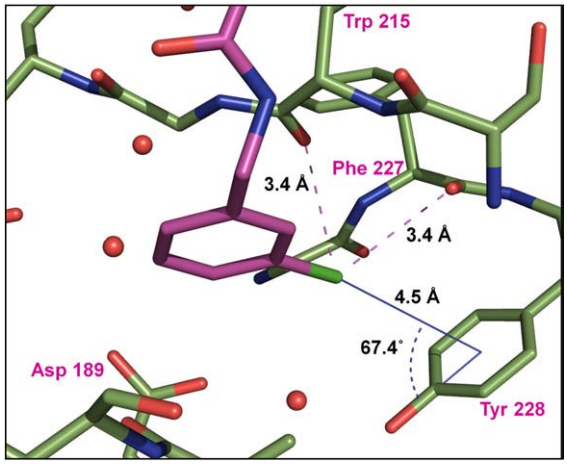

Fig. 4. Enlarged view of the local environment of the chlorine atom of inhibitor **1b**. The inhibitor is coloured magenta, and protein residues are coloured green, heteroatoms are colour-coded. Interactions are indicated by broken lines, and the chlorine–π interaction is indicated as a continuous blue line.

orientation, observed frequently for chlorine atoms pointing towards aromatic ring systems in a large dataset of small molecule crystal structures in the Cambridge Structural Database.[33] In terms of a recent evaluation of chlorine–π interactions in protein–ligand complexes[14], this orientation is characterized as the "edge on" geometry.

To obtain qualitative information about the residual mobility in the protein–ligand complex, special attention has been directed towards the crystallographically observed temperature factors. As their values are correlated significantly with the observed population parameters assigned to the hosted ligands, special care is needed in the interpretation. We analyzed the *B*-factors separately for each inhibitor relative to the neighbouring amino acids exposed to the binding site. This calibration was performed as an approximate normalization to correct for any superimposed quality difference in the diffraction properties of the crystals (Table 1). For **1b**, **3a** and **4,** the ratio of *B*-factors of ligand relative to the binding pocket residues is close to 1. This suggests equal thermal motion of the ligand and the binding cavity, and full occupancy of the inhibitors in the crystal.

This ratio increases significantly to 1.30 and 1.39 for inhibitors **1a** and **5,** respectively. One possible explanation for the enhanced *B*-factors of the ligands relative to the binding cavity is incomplete ligand occupancy. A second possible explanation is higher residual mobility of these ligands within the binding cavity due to a less firm arrest inside the S1 specificity pocket.

In order to avoid an unreasonable interpretation of occupancy differences of the inhibitors in the different crystal structures, the *B*-factors were compared in an alternative way. Each molecule was divided into the D-Phe-, Pro-and S1-occupant portion (Fig. 1) and the *B*-factors of these portions were evaluated separately.

Table 1. Binding data ΔG^0, ΔH^0 and $-T\Delta S^0$ (kJ/mol) of inhibitors as determined by ITC, sorted by ΔG^0, and evaluation of *B*-factors (Å^2) based on the corresponding crystal structures

Inhib.	ΔG^0 (kJ/mol)	ΔH^0 (kJ/mol)	$-T\Delta S^0$ (kJ/mol)	B Inhib.	B binding site	B Inhib./ B binding site	B D-Phe	B Pro	B D-Phe + Pro	B S1 occupant	B S1 occupant / B Pro	B S1 occupant / B D-Phe+ Pro
4	−46.1±0.6	−40.1±2.9	−6.1±3.4	11.4	10.9	**1.06**	12.3	11.6	11.9	10.5	**0.91**	**0.88**
6a	−38.4±0.2	−41.3±0.4	2.9±0.3									
6e	−37.3±0.3	−41.0±2.1	3.8±2.2									
6b	−37.2±0.5	−33.5±1.7	−3.7±2.2									
1c	−35.8±0.7	−34.5±0.4	−1.3±0.3									
1b	−35.4±0.8	−37.1±1.1	1.7±0.3	15.7	14.5	**1.08**	16.6	15.7	16.1	14.5	**0.92**	**0.90**
3a	−34.8±0.6	−28.5±0.7	−6.3±1.1	15.2	14.2	**1.06**	17.0	17.1	17.1	10.6	**0.62**	**0.62**
1d	−34.5±0.3	−38.0±1.1	3.5±0.8									
7a	−34.4±0.1	−31.9±1.3	−2.5±1.3									
1a	−31.3±0.2	−13.1±0.9	−18.2±0.7	18.8	14.5	**1.30**	17.3	18.0	17.7	21.3	**1.18**	**1.20**
3b	−32.9±0.5	−16.5±0.9	−16.4±0.4									
5	−31.7	−13.6	−18.1	26.2	18.9	**1.39**	26.2	24.9	25.5	28.2	**1.13**	**1.11**

Amino acids of the binding pocket (3 Å from the inhibitor): His57, Tyr60A, Pro60C, Trp60D, Glu97A, Asn98, Leu99, Asp102, Ile174, Asp189, Ala190, Cys191, Asp194, Ser195, Val213, Ser214, Trp215, Gly216, Glu217, Gly219, Cys220, Gly226, Phe227 and Tyr228. Errors indicate the standard deviation from at least triplicate measurements, for 5 only duplicate measurements could be recorded.

As only ligand portions are mutually compared, this comparison will not be affected by possible occupancy differences. The *B*-factors for the ligand atoms bound in the S1-pocket were compared to the D-Phe-Pro moiety of the molecule or to the Pro part alone (Table 1). Again, this ratio is lower and closer to 1 for **1b**, **3a** and **4**, indicating a more firm arrest of these ligands inside the S1 pocket. In contrast, for **1a** and **5**, the ligand portion inside the S1 pocket shows much higher *B*-factors than the remaining part of the molecule. Thus, ratios larger than 1 are observed. Combining both approaches for evaluating the *B*-factors, and considering the fact that **1a** and **5** have smaller P1 side chains than **1b**, **3a** and **4**, the conclusion can be made that the fluoro-and unsubstituted P1 side chains result in greater residual mobility for the ligand while bound in the thrombin active site.

Factorization into enthalpic and entropic contributions by isothermal titration calorimetry

In order to factorize the Gibbs free energy of binding ΔG^0, the standard binding enthalpy ΔH^0_{bind} and the dissociation constant K_D were determined by ITC for 12 inhibitors. The entropic contribution to binding $-T\Delta S^0$ has been calculated as difference between ΔG^0 (derived from K_D) and ΔH^0_{bind} according to the Gibbs–Helmholz equation. The results are given in Table 1. Despite their close similarity, the inhibitors show a large variation in the free energy of binding to thrombin, between –46 kJ/mol for the benzamidine **4** and –32 kJ/mol for the unsubstituted benzylamide **5**. As a general trend, the more potent inhibitors also experience a larger enthalpic contribution to binding. It is particularly noteworthy, that the benzylamides **1b-d,** substituted by a chlorine, bromine or iodine atom in the *meta* position, have a binding enthalpy almost as large as the salt-bridge-forming benzamidine **4**. This enthalpic effect is even larger for the *m*-chlorobenzylamides **6a** and **6e**, which have a second halogen atom at the opposing *para* position of the phenyl ring.

A significant drop in ΔH^0_{bind} is observed for **1a**, **3b** and **5**. Since thrombin accommodates the larger halogens, Br and I, without a significant drop in the ΔH^0_{bind}, and modelling experiments suggest that there is supposedly sufficient room in the S1 pocket to accommodate an ethyl group, the reduction ΔH^0_{bind} for **3b** is unlikely to be mainly due to steric reasons. The expected identical binding modes of **1a** and **5** were confirmed by crystallography. Their weakened enthalpic interaction is counterbalanced, at least in part, by an enhanced favourable entropic contribution to binding. This enthalpy–entropy compensation is believed to be a fundamental feature of weak non-covalent interactions and it is supported by our data. Interestingly, the thermodynamic profiles of **1a** and **5** do not differ, even though the unsubstituted **5** does not displace the conserved water molecule in the S1 pocket upon binding. The fluoro derivative **1a** must maintain an unoccupied void in this region. Previous analysis of similar cases have shown that capturing a

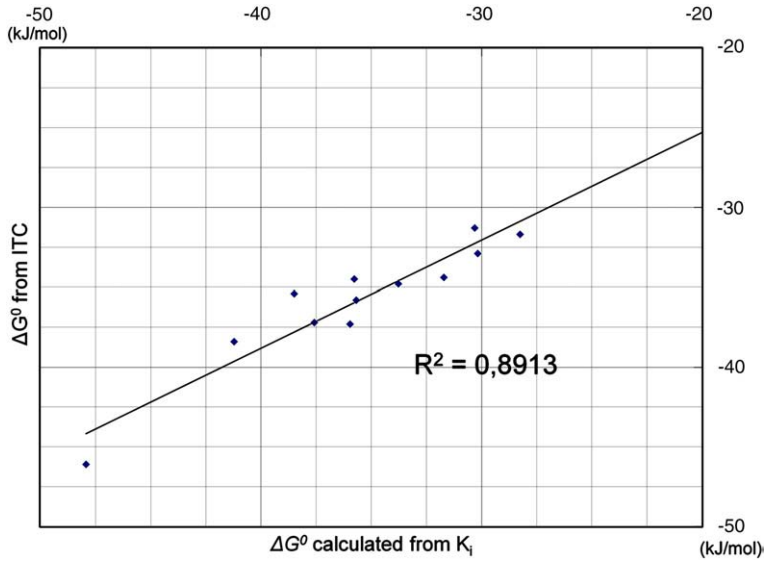

Fig. 5. Correlation between Gibbs free energies of binding ΔG as determined by ITC and as calculated from the kinetically determined K_i ($\Delta G = -RT \ln K_i$).

water molecule upon binding is entropically unfavourable and enthalpically beneficial for binding.[34] If the water molecule not displaced in the case of the binding of **5** to thrombin, superimposes an unfavourable entropic and beneficial enthalpic contribution, the thermodynamic signature of the latter ligand should be even less enthalpically driven and entropically disfavoured compared to **1a**. Most likely, several mutually compensating effects result in a virtually identical profile for the unsubstituted and fluoro derivative.

Figure 5 shows that for all 12 inhibitors the free energy of binding ΔG^0, as determined by ITC, exhibits a good correlation with the free energy for the binding event that can be derived from the inhibition constant:

$$\Delta G = -RT\ln K_i$$

as determined by the kinetic enzyme assay. Both methods to determine binding affinity substantially affirm each other, although the underlying principles and experimental conditions (e.g. displacement of the chromogenic substrate in the case of K_i) differ markedly.

Considerations with respect to differing desolvation energies

The binding affinity for the formation of a protein–inhibitor complex is strongly dependent on the desolvation energy that is required to desolvate polar groups of the interacting partners.[35] As very similar areas of the protein are buried upon ligand binding for the thrombin inhibitors studied here, we assume a constant contribution for the protein, whereas the desolvation energy of the inhibitors relates mainly to the solvation enthalpy in water. As the inhibitors investigated differ only in the substitution pattern at the phenyl ring occupying the S1-specificity pocket, a crude approximation roughly correlated with the expected desolvation free energy can be attempted by considering the experimentally determined water solubility of corresponding benzene derivates (Table 2).[36–38] These data show that substitution of the benzene ring with groups such as F, Cl, Br, I or methyl, is detrimental to water solubility. In contrast, the polar phenolic hydroxyl group enhances solubility significantly. Interestingly, substitution of benzene with a halogen atom leads to a significant decrease in solubility—and thus a likely decrease in desolvation energy—except for fluorine. This finding is further supported by our subsequent QM/MM study,[39] wherein the calculated desolvation energies of the larger halogen-containing inhibitors, **1b**, **1c**, and **1d** are significantly smaller than those for the fluorine inhibitor **1a**. The special properties of the methyl-containing inhibitor **3a** have also been investigated by QM/MM calculations and will be further addressed.[39] This behaviour seems to parallel the reduced binding enthalpy of **1a** and **5** (Table 1), indicating that a lower desolvation energy of the portion hosted in the S1 pocket seems to provide a crucial contribution to the enthalpy of binding. The approximate correlation of the measured binding enthalpy with the desolvation free energy of the benzylamides is emphasized by the fact that the inhibitors analogous to the least soluble benzene derivatives are generally the most enthalpic binders (Table 1). The highly soluble benzamidine derivative **4** is an exception to this observation, as its pronounced desolvation energy can be compensated

Table 2. Aqueous molar solubility of benzene derivatives as determined by *Yalkowsky et al.*, given on a logarithmic scale [36–38]

Derivative	Log solubility
1,4 Dichlorobenzene	–3.21
Iodobenzene	–2,87
Bromobenzene	–2.39
Chlorobenzene	–2.26
Toluene	–2.24
Fluorobenzene	–1.79
Benzene	–1.66
Phenol	–0.08

easily by the enthalpically favourable charged interaction with the deprotonated Asp189 at the bottom of the S1 pocket. This derivative still shows a favourable entropic contribution to binding even though the salt-bridge formed should be detrimental to the entropy of binding because it holds the ligand firmly in the S1 pocket, leading to reduced residual ligand entropy. However, the overall entropic benefit may result from an entropically favourable desolvation of the positively charged amidinium group, as this highly charged group will impose a strong local ordering effect on the surrounding water molecules in the bulk phase.

Correlation of thermodynamic and structural features

The detailed information collected for this series of inhibitors enables us to understand their differences with respect to binding energies, at least in qualitative terms. In this discussion, we assume that thermodynamically relevant changes within the host protein are virtually identical within the series of investigated ligands, and thus cancel out in a relative comparison of the inhibitors. Similarly, we assume the dynamics of the free ligands and thus their entropy before forming a complex to be virtually constant throughout the series, at least for the unpolar substituted benzylamides.

Due to the formation of a salt-bridge with the protein, the benzamidine derivative **4** binds with high binding enthalpy over-compensating the required high desolvation energy. Even though the compound is fixed rather firmly at the binding site, it still experiences a favourable entropic contribution to binding. This may result from the desolvation of the positively charged amidinium group (see above).

A major driving force for binding of the *meta*-substituted benzylamides **1b–d,** and **6a–e** is their low desolvation energy. In the case of the *m*-chloro compounds, an additional electrostatic component might be contributing to binding enthalpy. This chlorine–π attraction parallels cation–π interactions,[40–42] as it was demonstrated for similar inhibitors binding to the S1 pocket of factor Xa[9] and discussed recently for similar cases.[14] Although these compounds repel a tightly bound water molecule from the bottom of the S1 pocket in the uncomplexed state, their entropic contribution to binding is low. This observation agrees with the observation that the bound inhibitors exhibit low residual mobility in the binding pocket.

The unsubstituted inhibitor **5** and the slightly larger *m*-fluoro-substituted inhibitor **1a** have the attribute of low binding enthalpy in common. This correlates with their comparably high desolvation energy, re-emphasizing the importance of this entity, often neglected in affinity predictions. Another major difference relative to the *m*-halogen inhibitors is that both compounds lack a favourable electrostatic component to interact with the S1 binding pocket. Thus, they form a less firmly fixed complex with the protein, indicated by the enhanced residual mobility of the S1 portion of the ligands. Amplified residual or configurational entropy can be assigned to the ligands experiencing higher vibrational degrees of freedom in the complex.[43] The ligand internal loss of conformational entropy upon binding is assumed to be similar for all inhibitors of the series.

Conclusions

We have presented a congeneric series of thrombin inhibitors, for which the thermodynamic driving forces of complex formation have been studied in detail. The differing thermodynamic binding profiles of closely related ligands could be explained by individual differences in their potential to interact with solvent and protein. In particular, our data accentuate the strong influence of desolvation energy attributed to the part of the inhibitors that is buried deeply in the S1 pocket.

Furthermore, the series provides deeper insights into how spatial fixation affects the entropic contribution to binding. The examples show, that in general the enthalpically optimized ligands are the most potent binders, although obviously experiencing enthalpy–entropy compensation.

Special attention has been dedicated to the enthalpically favoured accommodation of the *m*-chloro-benzylamides in the S1 specificity pocket, which we explain by a electrostatically optimal positioning of the chlorine atom above the π electrons of Tyr228. This interaction of a chlorine substituent with the core of an aromatic amino acid resembles the well characterized cation–π interaction, and is often not considered by medicinal chemists as a possible strategy to gain binding affinity in ligand optimization.

Remarkably, the fluorine-substituted compound shows a very different thermodynamic signature compared to the other halogen derivates with respect to the investigated physiochemical properties. It seems that, due to the highly polar nature of the carbon-fluorine bond, the small fluorine atom interacts more favourably with water molecules in the bulk water than with the environment provided by the S1 pocket.

We hope that this study contributes to the fundamental understanding of molecular recognition of small molecules by a macromolecule and thus helps to transform concepts in drug design from predominantly phenomenological observations to more "rational" grounds driven by an understanding of all aspects involved.

Materials and Methods

Synthesis

Synthesis of Boc-D-Phe-L-Pro

L-Proline benzyl ester (772 mg, 3.77 mmol), Boc-D-phenylalanine (999 mg, 3.77 mmol), and *N*-hydroxybenzotriazole (560 mg, 4.14 mmol) were dissolved in 10 ml of

anhydrous dichloromethane (DCM). After 10 min 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (1.07 g, 8.29 mmol) was added, followed at 10 min later by the addition of diispropyl azodicarboxylate (8.29 mmol). The reaction mixture was stirred overnight. The DCM was removed and the residue was taken up in ethyl acetate, then washed three times with 1 M HCl, three times with saturated sodium bicarbonate, and three times with brine. The organic layer was dried with sodium sulfate, filtered, and concentrated, yielding 1.32 g of the product Boc-D-Phe-L-Pro benzyl ester as a colourless oil (78% yield). ^{1}H NMR ($C^2H_3O^2H$, 500MHz) 1.43 (s, 9H), 1.54–1.67(m, 1H), 1.75–2.02 (m, 3H), 2.80–3.02 (m, 3H), 3.32–3.35 (m, 2H), 4.23–4.39 (m, 1H), 4.44–4.51 (m, 1H), 5.14 (s, 2H), 7.24–7.38 (m, 10H). *m/v* (liquid chromatography-mass spectrometry, electrospray ionization): [M+Na]+, $C_{26}H_{32}N_2NaO_5$ requires 475.2, found 475.3.

Boc-D-Phe-L-Pro benzyl ester (633 mg, 1.40 mmol) was dissolved in 20 ml of methanol in a Parr hydrogenation bottle. The bottle was flushed with argon and 0.050 eq. of 10% Pd/C was added. The bottle was attached to a Parr hydrogenation apparatus and subjected to three charge/purge cycles with hydrogen gas. The reaction was then charged with hydrogen at 50–55 psi (1 psi ≈6.9 kPa) and shaken for 6 h. Liquid chromatography-mass spectrometry was used to monitor the disappearance of the starting material. When complete, the reaction mixture was filtered through celite to remove the Pd/C followed by several washes of the bottle and celite with methanol. The methanol fractions were combined and concentrated to give 405 mg of the product Boc-D-Phe-L-Pro as white solid (mp= 170–172 °C). ^{1}H NMR (C^2HCl_3, 500MHz): 1.25 (s, 9H), 1.68–1.96 (m, 3H), 2.12-2.27 (m, 1H), 2.52–2.62 (m, 1H), 2.86–2.98 (m, 1H), 3.09–3.16 (m, 1H), 3.5–4.59 (m, 1H), 4.34–4.35 (m, 1H), 4.62–4.69 (m, 1H), 5.31–5.39 (m, 1H), 7.21–7.35 (m, 5H). *m/v* (liquid chromatography-mass spectrometry, electrospray ionization): [M+Na]+, $C_{19}H_{26}N_2NaO_5$ requires 385.2, found 385.2.

General procedure for coupling Boc-D-Phe-L-Pro with benzyl amines

Boc-D-Phe-L-Pro (typically 1.0 mmol) was coupled with the benzyl amine hydrochlorides (1.5 eq.) corresponding to the final compound. After adding the Boc-D-Phe-L-Pro and the benzyl amine hydrochloride to the reaction flask, *N*-hydroxybenzotriazole (1.1 eq.) was added followed by 10 ml of anhydrous DCM. After all of the solids had dissolved completely, 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (1.1 eq.) was added followed by diispropyl azodicarboxylate (1.1 eq.). After stirring at room temperature for 12 h, the reaction was diluted with 20 ml of DCM and washed with 3 × 20 ml of 1 M HCl, 3 × 20 ml of saturated sodium bicarbonate, and 3 × 20 ml of brine. The organic layer was dried with sodium sulfate, filtered and concentrated to give the Boc-D-Phe-L-Pro-benzyl amide product. These crude products were recrystallized twice from methanol to give the purified products (30–80% yield). These Boc-protected peptides had sharp melting points, and gave ^{1}H NMR and mass spectrometry data confirming their structures.

General procedure for removing the Boc protecting group

The Boc-protected peptides (typically 0.14 mmol) were dissolved in 1.0 ml of anhydrous methanol and cooled in a bath of crushed ice. Gaseous HCl was bubbled through the reaction for 5 min and the reaction mixture was stirred for 1 h. The solvent was removed and the product was purified by semi-preparative HPLC (C18 RP-HPLC, 10% (v/v) methanol/90% (v/v) water for 10 min, then a linear gradient to 98% methanol over 45 min) to give the tested compounds shown in Fig. 1 in>95% purity. The ^{1}H NMR and mass spectrometry data confirmed the structure of each compound.

Bioassay

Kinetic inhibition of human thrombin (from Beriplast®, CSL Behring, Marburg, Germany) was determined photometrically at 405 nm using the chromogenic substrate Pefachrom tPa (LoxoGmbH, Dossenheim, Germany) as described[44] under the following conditions: 50 mM Tris–HCl, pH 7.4, 154 mM NaCl, 5 % DMSO, 0.1 % polyethylene glycol 8000 at 25°C using different concentrations of substrate and inhibitor. K_i ($n \geq 3$) were determined as described.[45]

Isothermal Titration Calorimetry

The ITC experiments were done with an MCS titration calorimeter (Wiseman et al. 1989) (Microcal, Inc., Northhampton, MA, USA).[54] Concentrations of inhibitor stock solutions in DMSO were determined by the weight of the corresponding hydrochlorides. Final concentration was achieved by diluting 1:40 (v/v) in the experimental buffer. Protein solution was prepared freshly for each experiment by dialysis of a thrombin sample (from Beriplast®, CSL Behring, Marburg, Germany) in the experimental buffer and final adjustment of the concentration of DMSO to 2.5% (v/v). The ITC measurements were routinely done at 25°C in 50 mM Tris–HCl, pH 7.8, 100 mM NaCl, 2,5% DMSO, 0.1% polyethylene glycol 8000. Inhibitor solutions (concentration in the range 0.2–0.5 mM, depending on the specific ligand) were degassed (~10 min) before use and titrated into the stirred cell (1.3513 ml) containing thrombin solution (0.01–0.02 mM) after a stable baseline had been achieved. The injection sequence consisted of an initial injection of 1.5 µl to prevent artefacts arising from the filling of the syringe (not used in data fitting), followed by injection of 7–10 µl each at intervals of 300 s until complete saturation of the enzyme binding sites was achieved. Raw data were collected and the area under each peak was integrated, followed by correction for heats of dilution and mixing by subtracting the final baseline consisting of small peaks of the same size to zero. Data were analysed using the ORIGIN Software (Microcal Inc.), by fitting a single-site-binding isotherm[54] that yields ΔH^0 (enthalpy of binding) and K_D (dissociation constant). Measurements were performed in triplicate; K was reproducible to within 10% and ΔH was reproducible to within 5%. Fig. 6 shows binding isotherms for the titration of **1a** (a), **1b** (b) and **6a** (c). The buffer dependence of ΔH^0 was tested for **4** in Tris, Hepes, tricine and pyrophosphate buffer. Despite different enthalpies of ionization for the buffers, only marginal differences in ΔH^0 were observed. Thus, any net proton exchange upon complex can be excluded or falls below 0.1 mol protons; in consequence, any protonation effect can be neglected in the discussion of binding energetics for all inhibitors.

Crystallisation and soaking

Human α-thrombin (from Beriplast®, CSL Behring, Marburg, Germany) was dissolved in crystallisation buffer (100 mM sodium phosphate, 350 mM NaCl, 10 mM benzamidine, pH 7.5) and after dialysis against the same buffer the sample was concentrated to 5 mg/ml. A 200 µl sample of this solution was mixed with 20 µl of an aqueous solution (500 mg/ml) of Hirugen (Bachem, Bubendorf,

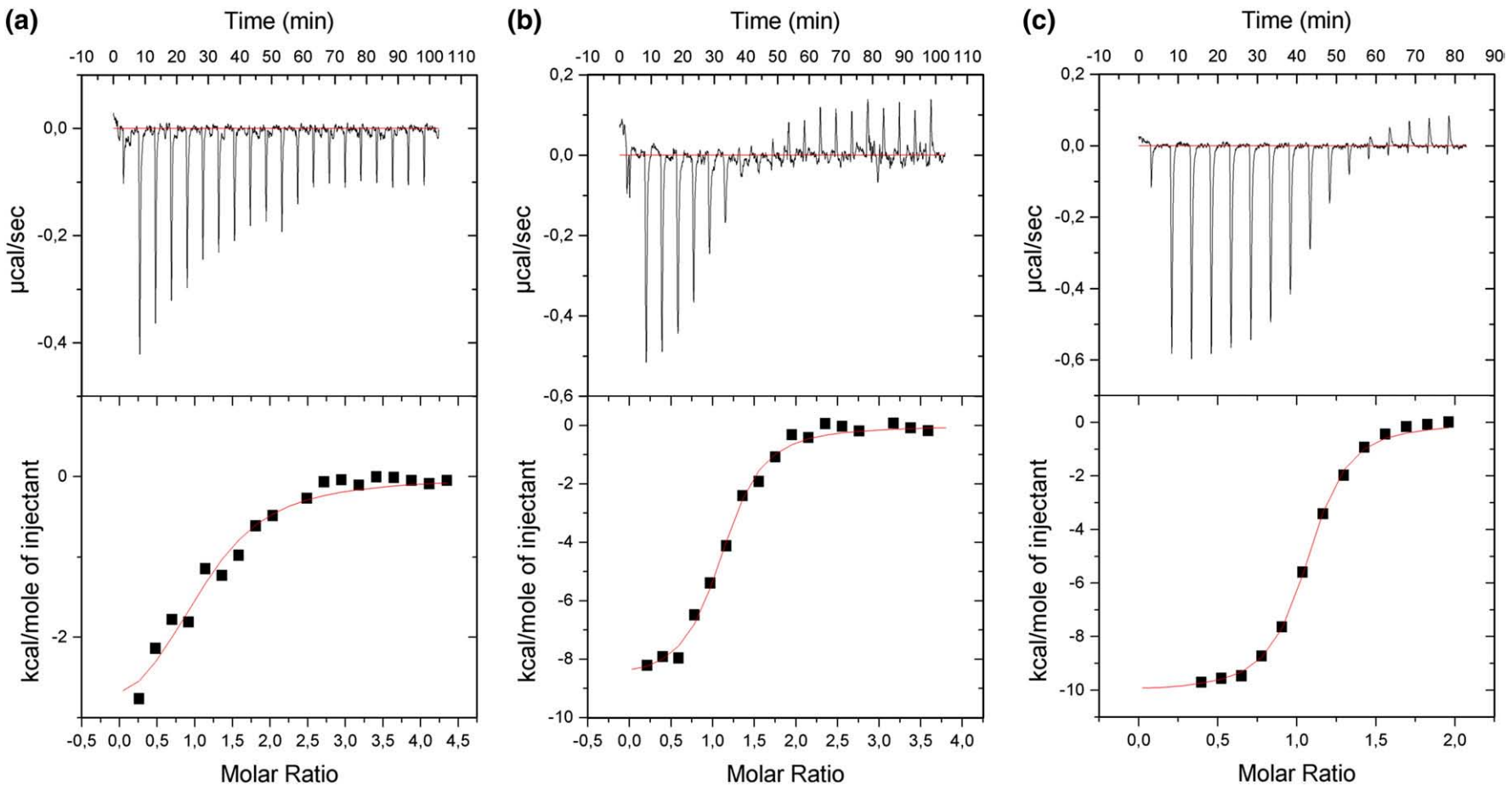

Fig. 6. Binding isotherms for the titration of **1a** (a), **1b** (b) and **6a** (c) into thrombin. The small peaks at the end of titration arise from the heat of dilution resulting from difficulties in perfectly matching the concentration of DMSO in the cell and in the syringe. This was corrected by subtracting the final baseline consisting of small peaks of the same size to zero.

Switzerland). After incubation for 10 h at 4 °C, crystallisation was carried out at 4° C by the vapour-diffusion, hanging-drop method using 28 % polyethylene glycol 8000, 100 mM sodium phosphate buffer (pH 7.5). After an equilibration period of 16 h, microseeding was done. For soaking, a mixture of 25% 10–40 mM solution of inhibitors in DMSO and 75% crystallisation buffer was prepared, in which medium-sized crystals without visible imperfections were soaked for 6–24 h. The uncomplexed structure resulted from five days soaking with a low-affinity inhibitor (K_i ~500 μM) in the crystallisation buffer, which was free of benzamidine. By this procedure, the crystals preincubated with benzamidine, which prevents self-digestion during crystallisation, were liberated from benzamidine in the S1 pocket, whereas the weak inhibitor did not occupy the active site in the amount necessary for detection in the scattering experiment.

Data collection and processing

Crystals were prepared for data collection at 110 K using a cryoprotectant solution of 20% (v/v) glycerol in crystallisation buffer. The datasets for **1a**, **1b**, **4, 5** and the *apo*-structure were collected with synchrotron radiation at BESSY beamlines 14.1 and 14.2 on a Marmosaic 225 mm CCD or Mar CCD 165 mm detector, respectively. Diffraction data for **3a** were collected at DESY beamline X13 on a Mar

Table 3. Data collection and refinement statistics for the *apo* structure and the five complex structures determined in this study

Structure (PDB entry)	THR-apo (3D49)	THR-4 complex (2ZDA)	THR-1b complex (2ZC9)	THR-1a complex (2ZDV)	THR-3a complex (2ZF0)	THR-5 complex (2ZFF)
A. *Data collection and processing*						
No. crystals used	1	1	1	1	1	1
Wavelength (Å)	0.91841	0.91841	0.97803	0.97803	0.80150	0.95373
Space group	*C*2	*C*2	*C*2	*C*2	*C*2	*C*2
Unit cell parameters						
a, *b*, *c* (Å)	70.4, 71.5, 72.5	70.3, 71.5, 72.4	70.4, 71.3, 72.9	70.5, 71.4, 72.8	70.0, 70.3, 72.5	70.1, 71.5, 72.5
β (°)	100.4	100.5	100.7	100.6	101.1	100.3
Matthews coefficient (Å^3/Da)	2.3	2.3	2.3	2.3	2.3	2.3
Solvent content (%)	47	46	47	47	46	46
B. *Diffraction data*[a]						
Resolution range (Å)	20 – 1.50 (1.53 – 1.50)	20 – 1.73 (1.76 – 1.73)	30 – 1.58 (1.61 – 1.58)	25 – 1.72 (1.78 – 1.72)	50 – 2.20 (2.25 – 2.20)	20 – 1.47 (1.50 – 1.47)
Unique reflections	54,430 (2419)	35,607 (1 692)	46,641 (1 725)	37,058 (3 536)	16,506 (819)	58,620 (2 837)
$R(I)_{sym}$ (%)	3.4 (17.6)	9.0 (25.2)	3.4 (23.5)	5.3 (16.9)	4.6 (16.6)	3.6 (49.9)
Completeness (%)	97.0 (86.8)	97.6 (93.1)	95.9 (71.1)	98.4 (94.8)	91.2 (69.1)	98.1 (96.7)
Redundancy	2.9 (2.7)	2.9 (2.2)	2.2 (1.7)	2.9 (2.3)	2.9 (2.5)	2.6 (2.6)
$I/\sigma(I)$	29.7 (4.9)	14.8 (3.3)	24.9 (3.4)	27.5 (3.6)	18.7 (6.1)	27.8 (2.6)
C. *Refinement*						
Resolution range (Å)	10 – 1.50	10 – 1.73	10 – 1.58	10 – 1.72	10 – 2.20	10 – 1.47
Reflections used in refinement (work/free)	53 163 / 2702	32 402 / 1 691	42 070 / 2197	33 538 / 1 776	15 099 / 775	52 843 / 2 814
Final *R* values for all reflections (work/free) (%)	20.2 / 28.5	17.2 / 22.8	18.0 / 23.3	18.7 / 24.3	17.0 / 26.7	14.6 / 20.3
Final *R* values for reflections with $F > 4\,\sigma\,F$ (work/free) (%)	19.8 / 23.7	16.9 / 21.4	17.1 / 22.0	17.9 / 23.0	16.1 / 25.2	14.0 / 19.3
Protein residues	250	252	251	251	252	252
Sodium ions	2	2	2	2	2	2
Inhibitor atoms	—	29	27	27	27	26
Water molecules	190	268	271	295	174	159
RMSD from ideality						
Bond lengths (Å)	0.011	0.008	0.010	0.008	0.005	0.010
Bond angles (°)	2.8	2.5	2.6	2.6	2.1	2.9
Ramachandran plot						
Residues in most favoured regions (%)	85.9	86.6	86.1	85.5	85.6	86.6
Residues in additionally allowed regions (%)	13.7	13.4	13.4	14.0	13.4	12.9
Residues in generously allowed regions (%)	0.4	—	0.5	0.5	1.0	0.5
Mean *B*-factor (Å^2)						
Protein	21.4	17.1	20.7	20.6	20.6	23.9
Binding site[b]	—	10.9	14.5	14.5	14.2	18.9
Inhibitor	—	11.4	15.7	18.8	15.1	26.2
Water molecules	28.4	26.7[c]	31.2	34.9	25.8	30.4[d]

[a] Numbers in parentheses are for the highest resolution shell.
[b] Amino acids of the binding site (3 Å from the inhibitor): His57, Tyr60A, Pro60C, Trp60D, Glu97A, Asn98, Leu99, Asp102, Ile174, Asp189, Ala190, Cys191, Asp194, Ser195, Val213, Ser214, Trp215, Gly216, Glu217, Gly219, Cys220, Gly226, Phe227, Tyr228.
[c] *B*-factor for additional water above Tyr228: 6.8 Å^2.
[d] *B*-factor for additional water above Tyr228: 35.9 Å^2.

CCD 165 mm detector. Data processing and scaling were performed using the HKL2000 package.[46]

Structure determination and refinement

The coordinates of human thrombin (PDB code 1H8D)[47] were used for initial rigid body refinement of the protein molecules followed by repeated cycles of maximum likelihood energy minimisation, simulated annealing and *B*-factor refinement using the CNS program package.[48] Refinement at later stages was done with the program SHELXL.[49] Here, at least 20 cycles of conjugate gradient minimisation were done with default restraints on bonding geometry and *B*-values. A randomly chosen 5% of all data were used for the calculation of R_{free} and were not used in the refinement. Amino acid side chains were fit into σ-weighted $2F_o - F_c$ and $F_o - F_c$ electron density maps using Coot.[50] After the first refinement cycle, water molecules and subsequently ions and ligand were located in the electron density and added to the model. Restraints were applied to bond lengths and angles, chiral volume, planarity of aromatic rings and van der Waals contacts. Multiple side chain conformations were built in case an appropriate electron density was observed and maintained during the refinement, and if the minor populated side chain showed at least 10% occupancy. During the last refinement cycles, riding H atoms were introduced without additional parameters. The final models were validated using PROCHECK.[51] Data collection, unit cell parameters and refinement statistics are given in Table 3. Analysis of temperature factors was done with Moleman,[52,53] distances and angles were measured in SYBYL 8.0 (TriposInc., USA). The figures were prepared using Isis Draw (MDL, San Leandro, USA) and Pymol 0.99†. The Protein Data Bank accession codes of the coordinates and structure factors of all X-ray structures are given in Table 3.

Protein Data Bank accession numbers

Coordinates and structure factors have been deposited in the Protein Data Bank with the following accession numbers: THR-**apo** 3D49; THR-**1a** complex 2ZDV; THR-**1b** complex 2ZC9; THR-**3a** complex 2ZF0; THR-**4** complex 2ZDA; and THR-**5** complex 2ZFF.

Acknowledgements

We kindly acknowledge CSL Behring, Marburg, for supplying us with generous amounts of human thrombin from the production of Beriplast®.

We thank the beamline support staff at DESY and BESSY for their advice during data collection, and the BMBF (support code 05ES3XBA/5) for generously supporting travel to BESSY/Berlin.

Author Contributions: MM, MZ and DH: The design and synthesis of the compounds as well as collaborative interpretation of the binding data. BB, CG, AH and GK: The testing of the compounds, x-ray determination of the binding mode, as well as collaborative interpretation of the binding data.

† http://pymol.sourceforge.net/

References

1. Steinmetzer, T. & Sturzebecher, J. (2004). Progress in the development of synthetic thrombin inhibitors as new orally active anticoagulants. *Curr. Med. Chem.* **11**, 2297–2321.
2. Danilewicz, J. C., Abel, S. M., Brown, A. D., Fish, P. V., Hawkeswood, E., Holland, S. J. *et al.* (2002). Design of selective thrombin inhibitors based on the (R)-Phe-Pro-Arg sequence. *J. Med. Chem.* **45**, 2432–2453.
3. Ho, S. J. & Brighton, T. A. (2006). Ximelagatran: direct thrombin inhibitor. *Vasc. Health Risk Manag.* **2**, 49–58.
4. Ambler, J., Bentley, D., Brown, L., Dunnet, K., Farr, D., Janus, D. *et al.* (1999). The discovery of orally available thrombin inhibitors: optimisation of the P1 pharmacophore. *Bioorg. Med. Chem. Lett.* **9**, 1103–1108.
5. Tucker, T. J., Brady, S. F., Lumma, W. C., Lewis, S. D., Gardell, S. J., Naylor-Olsen, A. M. *et al.* (1998). Design and synthesis of a series of potent and orally bioavailable noncovalent thrombin inhibitors that utilize nonbasic groups in the P1 position. *J. Med. Chem.* **41**, 3210–3219.
6. Lumma, W. C., Jr, Witherup, K. M., Tucker, T. J., Brady, S. F., Sisko, J. T., Naylor-Olsen, A. M. *et al.* (1998). Design of novel, potent, noncovalent inhibitors of thrombin with nonbasic P-1 substructures: rapid structure-activity studies by solid-phase synthesis. *J. Med. Chem.* **41**, 1011–1013.
7. Rittle, K. E., Barrow, J. C., Cutrona, K. J., Glass, K. L., Krueger, J. A., Kuo, L. C. *et al.* (2003). Unexpected enhancement of thrombin inhibitor potency with o-aminoalkylbenzylamides in the P1 position. *Bioorg. Med. Chem. Lett.* **13**, 3477–3482.
8. Young, M. B., Barrow, J. C., Glass, K. L., Lundell, G. F., Newton, C. L., Pellicore, J. M. *et al.* (2004). Discovery and evaluation of potent P1 aryl heterocycle-based thrombin inhibitors. *J. Med. Chem.* **47**, 2995–3008.
9. Chan, C., Borthwick, A. D., Brown, D., Burns-Kurtis, C. L., Campbell, M., Chaudry, L. *et al.* (2007). Factor Xa inhibitors: S1 binding interactions of a series of N-{(3S)-1-[(1S)-1-methyl-2-morpholin-4-yl-2-oxoethyl]-2-oxopyrrolidin-3-yl }sulfonamides. *J. Med. Chem.* **50**, 1546–1557.
10. Maignan, S., Guilloteau, J. P., Choi-Sledeski, Y. M., Becker, M. R., Ewing, W. R., Pauls, H. W. *et al.* (2003). Molecular structures of human factor Xa complexed with ketopiperazine inhibitors: preference for a neutral group in the S1 pocket. *J. Med. Chem.* **46**, 685–690.
11. Choi-Sledeski, Y. M., Kearney, R., Poli, G., Pauls, H., Gardner, C., Gong, Y. *et al.* (2003). Discovery of an orally efficacious inhibitor of coagulation factor Xa which incorporates a neutral P1 ligand. *J. Med. Chem.* **46**, 681–684.
12. Pinto, D. J., Orwat, M. J., Koch, S., Rossi, K. A., Alexander, R. S., Smallwood, A. *et al.* (2007). Discovery of 1-(4-methoxyphenyl)-7-oxo-6-(4-(2-oxopiperidin-1-yl)phenyl)-4,5,6,7-tetrahydro-1H-pyrazolo [3,4-c]pyridine-3-carboxamide (apixaban, BMS-562247), a highly potent, selective, efficacious, and orally bioavailable inhibitor of blood coagulation factor Xa. *J. Med. Chem.* **50**, 5339–5356.
13. Roehrig, S., Straub, A., Pohlmann, J., Lampe, T., Pernerstorfer, J., Schlemmer, K. H. *et al.* (2005). Discovery of the novel antithrombotic agent 5-chloro-N-({(5S)-2-oxo-3-[4-(3-oxomorpholin-4-yl)phenyl]-1,3-oxazolidin-5-yl}methyl)thiophene-2-carboxamide (BAY 59-7939): an oral, direct factor Xa inhibitor. *J. Med. Chem.* **48**, 5900–5908.

14. Imai, Y. N., Inoue, Y., Nakanishi, I. & Kitaura, K. (2008). Cl-pi interactions in protein-ligand complexes. *Protein Sci.* **17**, 1129–1137.
15. Sivaraman, S., Sullivan, T. J., Johnson, F., Novichenok, P., Cui, G., Simmerling, C. & Tonge, P. J. (2004). Inhibition of the bacterial enoyl reductase FabI by triclosan: a structure-reactivity analysis of FabI inhibition by triclosan analogues. *J. Med. Chem.* **47**, 509–518.
16. Tonge, P. J., Kisker, C. & Slayden, R. A. (2007). Development of modern InhA inhibitors to combat drug resistant strains of Mycobacterium tuberculosis. *Curr. Top. Med. Chem.* **7**, 489–498.
17. Sullivan, T. J., Truglio, J. J., Boyne, M. E., Novichenok, P., Zhang, X., Stratton, C. F. *et al.* (2006). High affinity InhA inhibitors with activity against drug-resistant strains of Mycobacterium tuberculosis. *ACS Chem. Biol.* **1**, 43–53.
18. Sivaraman, S., Zwahlen, J., Bell, A. F., Hedstrom, L. & Tonge, P. J. (2003). Structure-activity studies of the inhibition of FabI, the enoyl reductase from Escherichia coli, by triclosan: kinetic analysis of mutant FabIs. *Biochemistry*, **42**, 4406–4413.
19. Gohlke, H. & Klebe, G. (2002). Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew Chem. Int. Ed. Engl.* **41**, 2644–2676.
20. Lundqvist, T. (2005). The devil is still in the details-driving early drug discovery forward with biophysical experimental methods. *Curr. Opin. Drug Discov. Dev.* **8**, 513–519.
21. Klebe, G. (2006). Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov. Today*, **11**, 580–594.
22. Warren, G. L., Andrews, C. W., Capelli, A. M., Clarke, B., LaLonde, J., Lambert, M. H. *et al.* (2006). A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **49**, 5912–5931.
23. Whitesides, G. M. & Krishnamurthy, M. (2005). Designing ligands to bind proteins. *Quart. Rev. Biophys.* **38**, 385–395.
24. Talhout, R., Villa, A., Mark, A. E. & Engberts, J. B. (2003). Understanding binding affinity: a combined isothermal titration calorimetry/molecular dynamics study of the binding of a series of hydrophobically modified benzamidinium chloride inhibitors to trypsin. *J. Am. Chem. Soc.* **125**, 10570–10579.
25. Barratt, E., Bronowska, A., Vondrasek, J., Cerny, J., Bingham, R., Phillips, S. & Homans, S. W. (2006). Thermodynamic penalty arising from burial of a ligand polar group within a hydrophobic pocket of a protein receptor. *J. Mol. Biol.* **362**, 994–1003.
26. Barratt, E., Bingham, R. J., Warner, D. J., Laughton, C. A., Phillips, S. E. & Homans, S. W. (2005). Van der Waals interactions dominate ligand-protein association in a protein binding site occluded from solvent water. *J. Am. Chem. Soc.* **127**, 11827–11834.
27. Krishnamurthy, M., Bohall, B. R., Semetey, V. & Whitesides, G. M. (2006). The paradoxical thermodynamic basis for the interaction of ethylene glycol, glycine, and sarcosine chains with bovine carbonic anhydrase II: an unexpected manifestation of enthalpy/entropy compensation. *J. Am. Chem. Soc.* **128**, 5802–5812.
28. Mlinsek, G., Oblak, M., Hodoscek, M. & Solmajer, T. (2007). Thrombin inhibitors with novel P1 binding pocket functionality: free energy of binding analysis. *J. Mol. Model.* **13**, 247–254.
29. Gerlach, C., Smolinski, M., Steuber, H., Sotriffer, C. A., Heine, A., Hangauer, D. G. & Klebe, G. (2007). Thermodynamic inhibition profile of a cyclopentyl and a cyclohexyl derivative towards thrombin: the same but for different reasons. *Angew Chem. Int. Ed. Engl.* **46**, 8511–8514.
30. Bode, W. (2005). The structure of thrombin, a chameleon-like proteinase. *J. Thromb. Haemost.* **3**, 2379–2388.
31. Gerlach, C., Munzel, M., Baum, B., Gerber, H. D., Craan, T., Diederich, W. E. & Klebe, G. (2007). KNOBLE: a knowledge-based approach for the design and synthesis of readily accessible small-molecule chemical probes to test protein binding. *Angew Chem. Int. Ed. Engl.* **46**, 9105–9109.
32. Stubbs, M. T., Oschkinat, H., Mayr, I., Huber, R., Angliker, H., Stone, S. R. & Bode, W. (1992). The interaction of thrombin with fibrinogen. A structural basis for its specificity. *Eur. J. Biochem.* **206**, 187–195.
33. Swierczynski, D., Luboradzki, R., Dolgonos, G., Lipkowski, J. & Scheider, H. (2005). Non-covalent interactions of organic halogen compounds with aromatic systems- analyses of crystal structure data. *Eur. J. Org. Chem.* **6**, 1172–1177.
34. Petrova, T., Steuber, H., Hazemann, I., Cousido-Siah, A., Mitschler, A., Chung, R. *et al.* (2005). Factorizing selectivity determinants of inhibitor binding toward aldose and aldehyde reductases: structural and thermodynamic properties of the aldose reductase mutant Leu300Pro-fidarestat complex. *J. Med. Chem.* **48**, 5659–5665.
35. Shimokhina, N., Bronowska, A. & Homans, S. W. (2006). Contribution of ligand desolvation to binding thermodynamics in a ligand-protein interaction. *Angew Chem. Int. Ed. Engl.* **45**, 6374–6376.
36. Valvani, S. C., Yalkowsky, S. H. & Roseman, T. J. (1981). Solubility and partitioning IV: Aqueous solubility and octanol-water partition coefficients of liquid nonelectrolytes. *J. Pharm. Sci.* **70**, 502–507.
37. Yalkowsky, S. H. & Valvani, S. C. (1980). Solubility and partitioning I: Solubility of nonelectrolytes in water. *J. Pharm. Sci.* **69**, 912–922.
38. Bodor, N. & Huang, M. J. (1992). A new method for the estimation of the aqueous solubility of organic compounds. *J. Pharm. Sci.* **81**, 954–960.
39. Freindorf, M. (2009). Protein ligand binding affinity: a detailed investigation into the surprising potency provided by a meta-chlorobenzyl side chain in the thrombin S1 pocket. In the Press.
40. Gallivan, J. P. & Dougherty, D. A. (1999). Cation-pi interactions in structural biology. *Proc. Natl Acad. Sci. USA*, **96**, 9459–9464.
41. Mecozzi, S., West, A. P., Jr & Dougherty, D. A. (1996). Cation-pi interactions in aromatics of biological and medicinal interest: electrostatic potential surfaces as a useful qualitative guide. *Proc. Natl Acad. Sci. USA*, **93**, 10566–10571.
42. Meyer, E. A., Castellano, R. K. & Diederich, F. (2003). Interactions with aromatic rings in chemical and biological recognition. *Angew Chem. Int. Ed. Engl.* **42**, 1210–1250.
43. Chang, C. E., Chen, W. & Gilson, M. K. (2007). Ligand configurational entropy and protein binding. *Proc. Natl Acad. Sci. USA*, **104**, 1534–1539.
44. Sturzebecher, J., Sturzebecher, U., Vieweg, H., Wagner, G., Hauptmann, J. & Markwardt, F. (1989). Synthetic inhibitors of bovine factor Xa and thrombin comparison of their anticoagulant efficiency. *Thromb. Res.* **54**, 245–252.
45. Dixon, M. (1972). The graphical determination of K m and K i. *Biochem. J.* **129**, 197–202.

46. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
47. Skordalakes, E., Dodson, G. G., Green, D. S., Goodwin, C. A., Scully, M. F., Hudson, H. R. *et al.* (2001). Inhibition of human alpha-thrombin by a phosphonate tripeptide proceeds via a metastable pentacoordinated phosphorus intermediate. *J. Mol. Biol.* **311**, 549–555.
48. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.
49. Sheldrick, G. M. & Schneider, T. R. (1997). SHELXL: high-resolution refinement. *Macromol. Crystallogr. B*, **277**, 319–343.
50. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D*, **60**, 2126–2132.
51. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
52. Kleywegt, G. J., Zou, J. Y., Kjeldgaard, M. & Jones, T. A. (2001). Around O. In *International Tables for Crystallography* (Rossmann, M. G. & Arnold, E., eds), vol. F, pp. 353–356, Kluwer Academic Publishers, Dordrecht.
53. Kleywegt, G. J., Zou, J. Y., Kjeldgaard, M. & Jones, T. A. (2001). Around O. In *International Tables for Crystallography* (Rossmann, M. G. & Arnold, E., eds), vol. F, pp. 366–367, Kluwer Academic Publishers, Dordrecht.
54. Wiseman, T., Willistou, S. & Brandts, S. F. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **17**, 131–137.