Compounds Binding to the S2−S3 Pockets of Thrombin

Mikael Nilsson,*,† Markku Hämäläinen,‡ Maria Ivarsson,‡ Johan Gottfries,§ Yafeng Xue,§ Sebastian Hansson,§ Roland Isaksson,*,† and Tomas Fex*,§

School of Pure and Applied Natural Sciences, University of Kalmar, S-391 82 Kalmar, Sweden, GE Healthcare Bio-Sciences AB, S-754 50 Uppsala, Sweden, R&D AstraZeneca, S-431 83 Mölndal, Sweden

Received September 19, 2008

A set of compounds designed to bind to the S2−S3 pockets of thrombin was prepared. These compounds included examples with no interactions in the S1 pocket. Proline, a common P2 in many thrombin inhibitors, was combined with known P3 residues and P1 substituents of varying size and lipophilicity. Binding constants were determined using surface plasmon resonance (SPR) biosensor technology and were found to be in good agreement with results from an enzyme assay. A dramatic increase in affinity (100−1000 times) was seen for compounds incorporating an amino group capable of forming a hydrogen bond with gly216 in the protein backbone. The ligand efficiency was increased by including substituents that form stronger hydrophobic interactions with the P1 pocket. The binding mode was confirmed by X-ray analysis, which revealed the anticipated binding motif that included hydrogen bonds as well as a tightly bound water molecule. A QSAR model indicated that hydrogen bonding and lipophilicity were important for the prediction of binding constants. The results described here may have implications for how directed compound libraries for shallow protein pockets, like S2 and S3 in serine proteases, can be designed.

Introduction

Proteases represent an important target class, with several pharmacologically relevant targets, but it has been notoriously difficult to find useful inhibitors.[1−4] The inhibitors are often peptide-like and try to mimic the amino acids that are characteristic for the substrate. Serine protease inhibitors can usually be viewed as consisting of three units, P1, P2, and P3. P1 binds in the S1 pocket, which is usually quite deep, whereas P2 and P3 bind in the much more shallow S2 and S3 pockets.

Thrombin is a serine protease that is critically involved in blood coagulation by converting soluble fibrinogen into insoluble fibrin. The Asp189 in the S1 pocket of thrombin binds strongly to basic P1 residues, especially arginine and arginine-mimics such as benzamidine. Figure 1 shows the X-ray structure of the thrombin inhibitor melagatran[5] in complex with thrombin, also detailing important hydrogen bonds between the inhibitor and the protein backbone (PDB IK22).

In the present study, we have prepared compounds having a proline P2 scaffold and P3 residues from known thrombin inhibitors in combination with P1 substituents of varying size and lipophilicity. This enabled the importance of P1 for binding to the enzyme to be investigated and to assess whether compounds without P1 could be detected as binders. Also, the importance of P3 substituents with capacity to form hydrogen bonds to the protein backbone was studied. Using SPR biosensor measurement of binding affinities, we found that an amino group in P3 had a remarkably strong effect on potency and an X-ray structure confirmed its paticipation in a hydrogen bonding network. A significant increase in ligand efficiency was seen when lipophilic P1 substituents penetrated deeper into the P1 pocket.

Results and Discussion

The compounds making up the small directed library are shown in Table 1, along with their corresponding affinity and IC_{50} values. Proline is a common P2 residue in many thrombin inhibitors, and thus the potential to form two important hydrogen bonds (similar as depicted for melagatran in Figure 1) was retained. Lipophilic D-amino acids were incorporated in the P3 position. In addition, building blocks without the amino group were used, and the amino group was also replaced with a hydroxyl group. Binding constants of up to 1 mM could be determined using SPR biosensor technology. Thrombin inhibition was also tested in a traditional enzyme assay using a maximum compound concentration of 167 μM. For some compounds, extrapolation of inhibition curves allowed determination of approximate IC_{50} values even above this concentration. For those compounds which yielded IC_{50} values in the enzyme assay, it can be seen that there is good correlation between the two methods (Figure 2).

An overview of some interesting structure−activity relationships is shown in Figure 3.

When R1 is a methyl group, only compound **25** (K_D 38 μM, Table 1)) with a very lipophilic P3 had high enough affinity to be readily detected. Compounds **21** (K_D 550 μM) and **29** (K_D 790 μM) are much less potent and the affinities are barely detectable. Looking at compounds **14** (K_D 99 μM), **22** (K_D 45 μM), **26** (K_D 3.1 μM), and **30** (K_D 27 μM), it can be seen that replacing methyl with *n*-propyl increases potency by more than 10 times. This correlates with increased lipophilicity of the substituent and shows that hydrophobic interactions in the S1 region of the enzyme are favorable.

The methoxyethyl substituent in compounds **15** (K_D 750 μM), **23** (K_D 140 μM), **27** (K_D 6.8 μM), and **31** (K_D 110 μM) is more polar and results in compounds with intermediate potency. The thienylmethyl substituent, in compounds **4** (K_D 700 μM), **8**

* To whom correspondence should be addressed. For T.F.: . phone, +46 31 7065636; fax, +46 31 7763818; E-mail, Tomas.Fex@astrazeneca.com. For R.I.: phone, +46480446751; fax, +46480447305; E-mail, Roland.Isaksson@hik.se. For M.N.: phone, +46480446480; fax, +46480447305; E-mail, Mikael.Nilsson@hik.se. Present address for M.N.: Cambrex Karlskoga AB, SE-69185 Karlskoga Sweden; phone, +46586783116; fax, +46586783129; E-mail, mikael.nilsson@cambrex.com.

† School of Pure and Applied Natural Sciences, University of Kalmar.

‡ GE Healthcare Bio-Sciences AB.

§ R&D AstraZeneca.

10.1021/jm8011849 CCC: $40.75
Published on Web 04/16/2009

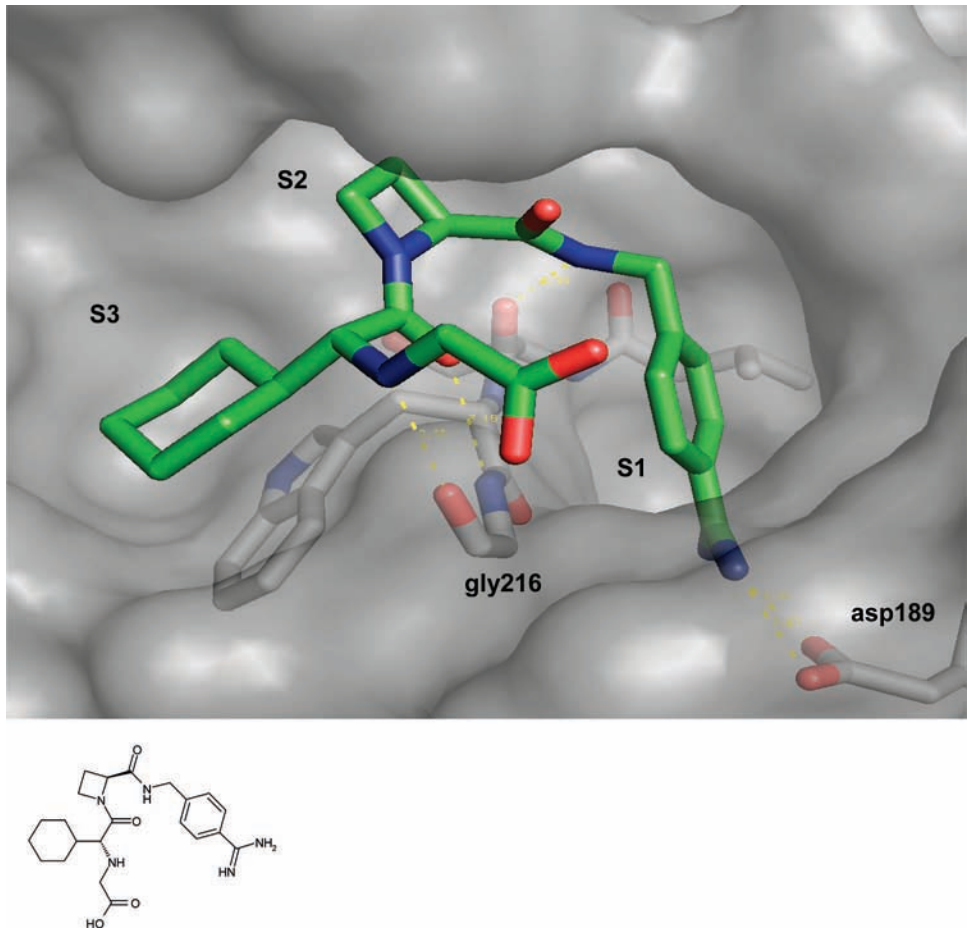


Figure 1. Melagatran in the active site of thrombin with the H-bonds to the enzyme backbone highlighted (PDB 1K22).

(K_D 34 μM), **12** (K_D 1000 μM), **16** (K_D 23 μM), **20** (K_D 250 μM), **24** (K_D 2.2 μM), **28** (K_D 0.27 μM), and **32** (K_D 1.1 μM) clearly produces the most potent compounds, with activity observed for all different R2 substituents, underlining that large contributions of binding energy can be obtained through interactions in the S1 pocket. Compound **33** (K_D 49 μM), with an isobutyl substituent is less active than the corresponding *n*-propyl derivative **30** (K_D 27 μM), indicating that the extra methyl group may result in unfavorable steric interactions. Compound **34** (K_D 9 μM) is the related cyclohexylmethyl derivative, but here the slight increase in potency does not parallel the substantial increase in lipophilicity as compared to the propyl substituent. This is rather surprising because previous investigations have shown that the cyclohexylamine analogue is a potent S1 binder.[5]

The results for different R2 substituents are striking. Those having an amino group are by far the most potent, with 100−1000 times higher affinities than their nonamino equivalents. This is seen when comparing compounds **24** (K_D 2.2 μM), **28** (K_D 0.27 μM), and **32** (K_D 1.1 μM) with **4** (K_D 700 μM), **8** (K_D 34 μM), and **12** (K_D 1000 μM). The amino group can form a hydrogen bond to gly216 in the enzyme (see Figure 1), which contributes very significantly to the total binding energy. A hydroxyl group seems to be much less efficient in this respect, however, as seen when comparing **36** (no binding) and **31** (K_D 110 μM). Even if compound **16** (K_D 23 μM) can not be compared with **28** (K_D 0.27 μM) in a straightforward manner, there is also an indication here that the amino group is preferred over the hydroxyl.

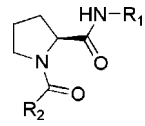
Compound **35** (K_D 35 μM) is a phenylmethylsulfonamide derivative of **29** (K_D 790 μM), and this added substituent increases potency by about 20 times. One of the sulfonyl oxygens can form a hydrogen bond to gly218, while the phenyl group can reach toward the so-called S1β binding site above the S1 pocket.[6−8]

As can be seen in Figure 4, those compounds where R2 contains an amino functionality exhibit higher ligand efficiency than those without. Ligand efficiency (LE) is defined as $-\Delta G$/number of non-hydrogen atoms in the molecule. LE are obtained from SPR biosensor measurements. This suggests that for optimizing a hit (i.e., confirmed binder), finding additional hydrogen bonding possibilities may be very productive in terms of increasing potency while minimizing the molecular weight. Another interesting observation is that the ligand efficiency increases almost linearly with an increase in size of the P1 residue. This finding is somewhat unusual compared to what is often seen during optimizations, where constant or decreasing ligand efficiency is more common.[9] The S2−S3 pockets are relatively shallow, and compounds binding in them will have relatively weak affinity compared to compounds binding in the S1 pocket. Thus, ligand efficiency will increase when compounds can also utilize the S1 pocket for binding.

For a more detailed analysis, K_D values were used for QSAR[a] calculations. The library compounds were characterized by chemical descriptors relating to lipophilicity (ClogP, ClogD, and molecular refractivity MR), size and shape (MW and Volume), flexibility and rigidity (rotatable bonds and ring count), polar features (H-bond donors and acceptors and polar surface

[a] Abbreviations: SPR, surface plasmon resonance; HTS, high throughput screening; DMPK, distribution metabolism and pharmacokinetics; QSAR, quantatative structure activity relationships; OPLS, orthogonal projections to latent structures.

Table 1. Synthesized Compounds Including K_D and IC_{50} Data[a]



Nr	R1	R2	K_D (IC_{50}) μM	Nr	R1	R2	K_D (IC_{50}) μM
1	CH_3		NB (NA)	21	CH_3		550 (NA)
2	nC_3H_7		NB (NA)	22	nC_3H_7		45 (210)
3	$CH_2CH_2OCH_3$		620 (NA)	23	$CH_2CH_2OCH_3$		140 (NA)
4			700 (NA)	24			2.2 (13)
5	CH_3		NB (NA)	25	CH_3		38 (230)
6	nC_3H_7		270 (NA)	26	nC_3H_7		3.1 (18)
7	$CH_2CH_2OCH_3$		NB (NA)	27	$CH_2CH_2OCH_3$		6.8 (89)
8			34 (120)	28			0.27 (1.6)
9	CH_3		NB (NA)	29	CH_3		790 (NA)
10	nC_3H_7		NB (NA)	30	nC_3H_7		27 (120)
11	$CH_2CH_2OCH_3$		NB (NA)	31	$CH_2CH_2OCH_3$		110 (NA)
12			1000 (NA)	32			1.1 (4.2)
13	CH_3		NB (NA)	33	$CH_3CH(CH_3)_2$		49 (230)
14	nC_3H_7		99 (380)				
15	$CH_2CH_2OCH_3$		750 (NA)	34			9 (20)
16			23 (49)				
17	CH_3		NB (NA)	35	CH_3		35(170)
18	nC_3H_7		NB (NA)				
19	$CH_2CH_2OCH_3$		NB (NA)				
20			250 (NA)	36	$CH_2CH_2OCH_3$		NB (NA)

[a] NB = BIACORE, K_D > 1 mM, (NA) = enzyme assay, IC_{50} > 400 μM.

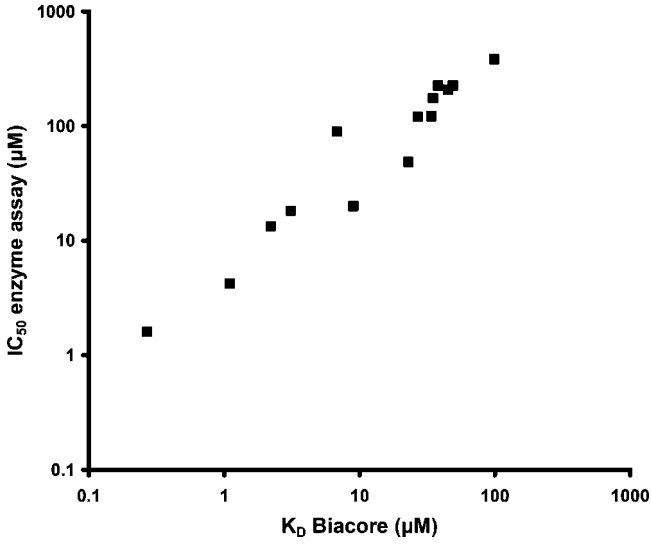


Figure 2. The correlation between IC_{50} and K_D.

area), and pK_a (indicator of base and, if available, the actual pK_a was calculated). The descriptors that were generated (k = 15) were correlated to the K_D values. Compounds generating IC_{50} and K_D values were included, n = 24. A QSAR model was obtained, which by OPLS[10] provided a 3 OPLS component model (one predictive and two orthogonal PLS components) with 78% explained Y variance (R^2 = 0.78 and Q^2 = 0.73). The prediction power of the QSAR model, as estimated by cross-validation, provided an RMSEE of 0.49 pK_D units, and the observed versus predicted values are plotted in Figure 5. Interpretation of model coefficients indicated that H-bonding and lipophilicity were of importance for prediction (Figure 6). First the presence of H-bond donor(s), but not acceptors, appeared essential, although the pK_a range included (7.2−8.4 by ACD laboratories) revealed only minor difference between the affinities. Second, lipophilicity appeared important for the QSAR model, indicated by significant ClogP and MR QSAR loadings (i.e., OPLS p_{1p}, significance tested by jack-knifing). The LogD descriptors were close to nonsignificance, which in combination with the H-bond donor information supports the conclusion that the presence of a H-bond donor, rather than its basicity, generated the binding energy. The contribution to affinity of the R1 substituents fits well with the

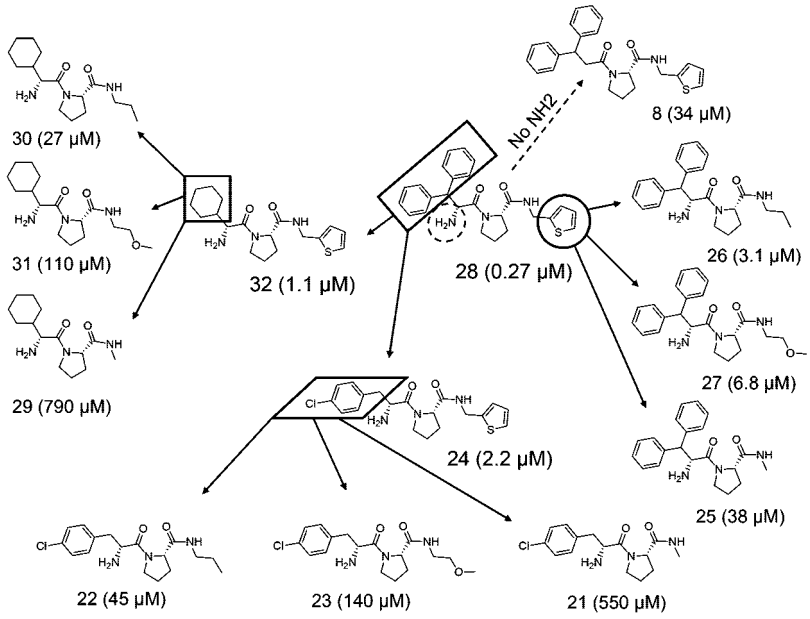


Figure 3. SAR for selected compounds showing the effect on affinities (K_D) of variation in R1 (methyl, *n*-propyl, methoxyethyl, thienylmethyl) for three different R3 substituents.

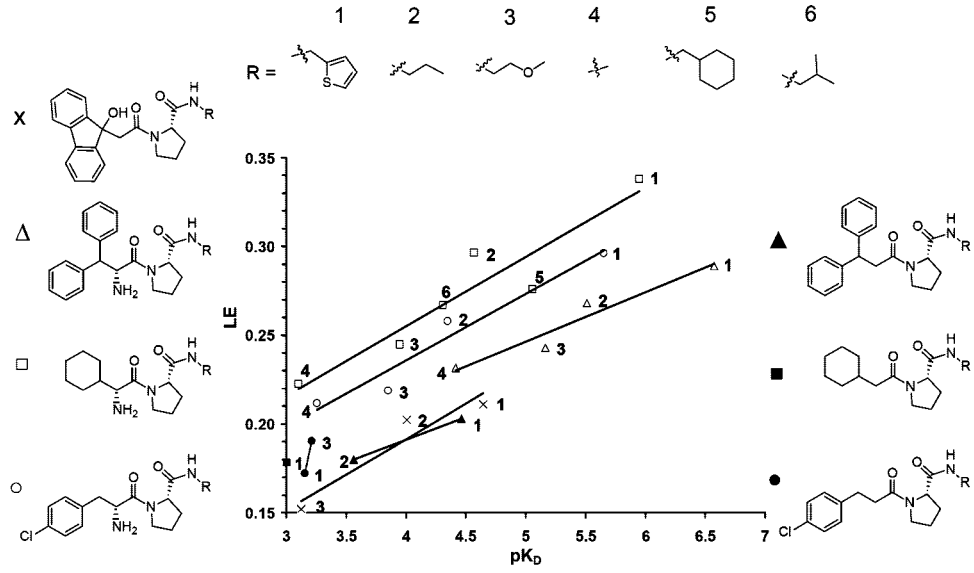


Figure 4. Ligand efficiency (LE) as a function of pK_D for selected series. LE = ($-\Delta G/N$); N is the number of non-hydrogen atoms, and $\Delta G = -$ RT $\ln(1/K_D)$.

QSAR model comprising positive and significant coefficients for ClogP and MR.

The quite dramatic effect observed for the amino group may have some important implications. Molecules that can form three interactions have three points of attachment to the enzyme, which may result in more rigid and stable complexes. In light of the fact that the hydrogen bond is formed between the amino group and a carbonyl group on the surface of the enzyme, the increased binding is even more intriguing. Thus, when a binder has been identified, one option for improving potency should be to look for potential additional hydrogen bonding interactions. If found, these can provide dramatic increases in potency. This is somewhat contradictory to the general notion that increased potency during lead optimization is achieved by increasing bulk and lipophilicity.

Compound **26**, which showed quite good potency and a P1 that could occupy only part of the S1 pocket, was selected for crystallization with thrombin (Figures 7−8). The X-ray structure obtained was in good agreement with the anticipated binding mode. There are three strong hydrogen bonds (all around 2.9 Å) formed between the inhibitor and thrombin. These hydrogen bonds contribute significantly to the observed binding affinity. In addition, the propyl tail of the inhibitor fits into the entrance of S1 and the diphenyl group fills into the S3 pocket. Together with the pyrrolidine group, which occupies the hydrophobic space formed by Tyr83/Trp86/Leu132 in S2, the shape of the inhibitor matches/fills the binding pockets of the protein very well. Furthermore, there are a number of water molecules surrounding the inhibitor, some of which bridge the interaction between the inhibitor and the protein. In particular, Wat154 seems to stabilize the inhibitor conformation by having hydrogen bonds to both the amino group and the amide carbonyl group. This may contribute to the overall high affinity of compounds with a terminal amino group.

The results from this study show that synthesis of P2−P3 libraries may be a useful approach to find novel serine protease

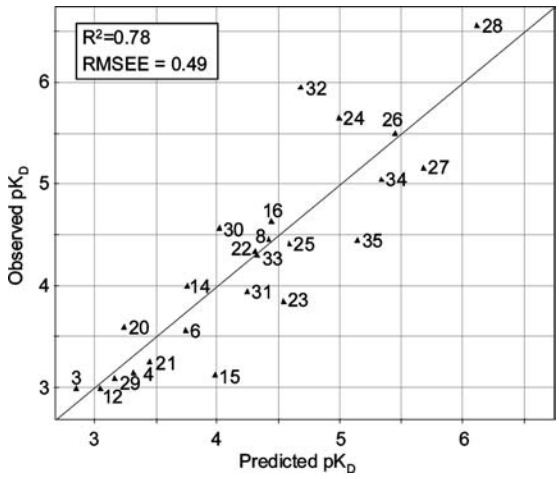


Figure 5. Observed pK_D values for compounds (numbering according to Table 1) binding to thrombin assessed by SPR plotted against the predicted pK_D. The predictions were generated by cross validation ($Q^2 = 0.71$), with an RMSEE = 0.49, using an OPLS model explaining 78% of pK_D according to R^2.

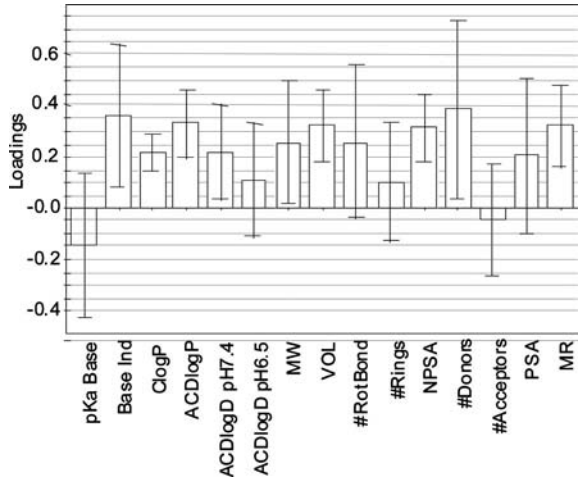


Figure 6. QSAR coefficients (i.e., P_{1p} loadings, confidence intervals by jack-knifing), using OPLS for regression with compounds binding to thrombin (pK_D) from SPR assessments as Y ($R^2 = 0.78$; $Q^2 = 0.73$; RMSEE = 0.49).

inhibitors. The importance of hydrogen bonding interactions is underlined, and compounds should be capable of accepting and/or donating hydrogen bonds. Once a binder has been identified, a further increase in potency can be achieved by a combination of adding new hydrogen bonding groups and changes that increase hydrophobicity. Hydrogen bonds to the protein surface can greatly enhance binding.

Protein−protein interactions are often made up of interactions between protein surfaces with relatively shallow pockets, perhaps resembling the shallow S2 and S3 pockets in proteases. Thus, the results presented here may have implications for how we try to find molecules that can interact with proteases and hence to develop effective drugs against these very challenging targets.

Experimental Section

SPR Methodology. All experiments where performed with a Biacore A100 instrument using Series S sensor chip CM5 (GE Healthcare Bio-Sciences AB, Uppsala, Sweden), 10 mM phosphate buffer with 137 mM NaCl, 2.7 mM KCl, 0.05% surfactant P20 (GE Healthcare Bio-Sciences AB, Uppsala, Sweden), and 5% dimethyl sulfoxide (Riedel-de Haën, Seelze, Germany) was used as running buffer.

The proteins where amine coupled (EDC/NHS, amine coupling kit, GE Healthcare Bio-Sciences AB, Uppsala, Sweden) to the sensor chip using a 10 mM NaAc pH 5.0 as a coupling buffer (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Typical immobilization levels were 5000−6000 RU (resonance units) for wt-thrombin (human α-thrombin, Haematologic Technologies Inc., VT) and blocked thrombin (human α-thrombin-DFP, Haematologic Technologies Inc., VT). Routinely, the following concentrations (μg/mL) and injection times (min) were used for immobilizations: Thrombin, 20 μg/mL, 5 min, and Thrombin-DFP, 20 μg/mL, 7 min. The flow rate was 30 μL/min. Test substances and positive controls were injected in a cycle with a contact time of 30 s and dissociation time of 30 s, followed by an extra wash of the flow system with 50% DMSO. The effect of bulk refractive index variations due to potential slight variations in DMSO content of the samples was compensated using a standard calibration routine implemented in the A100 Evaluation software. The calibration was performed in the 4.5−5.8% DMSO range.

The first affinity screen was run using a concentration range of 3−200 μM. Affinity calculations for the medium affinity compounds (K_D 6−100 μM) were based on this concentration range. High affinity compounds (0.3−4 μM) were analyzed using a concentration range of 1 nM to 10 μM. Nonbinders and weak binders were run using a concentration range of 30−500 μM to 7−1000 μM. Affinity calculations for the weak binders were based on these higher concentration ranges. The compounds were injected in a series of increasing concentrations. The response used for affinity calculations was the differential steady-state binding signal between wt-thrombin and the blocked thrombin, i.e., the wt-thrombin signal minus that of the blocked thrombin. For equilibrium binding responses of compounds binding to thrombin (wt minus blocked), the affinity (K_D) was calculated by fitting to the equation of a single-site binding model ($R = (\text{conc} \times R_{max})/(\text{conc} + K_D)$. A minimum of three replicate measurements were made for each compound. The average relative standard deviation of the affinity estimates was 20.7%.

Enzyme Assay. The thrombin inhibitor potency was measured with a chromogenic substrate method in a robotic microplate processor, using 96-well, half-volume microtiter plates.[11] The linear absorbance increase values were determined by measurements at 405 nm (37 °C) during 40 min with Melagatran as control substance. The IC_{50} values were calculated by fitting the data to a three-parameter equation by Microsoft XLfit. For some compounds with $IC_{50} > 167$ μM, an approximate IC_{50} value was obtained by extrapolation of inhibition curves. For melagatran, the $IC_{50} = 0.008$ μM.

Experimental Procedure Crystallization and X-Ray Structure Determination. Crystallization and soaking of compound: human α-thrombin (Factor IIa) was purchased from Enzyme Research Laboratories. Preformed thrombin−hirudin complex (stored at 4 °C) was used for crystallization. Apo-crystals were grown using hanging-drop method with microseeding. The drop was made by mixing 1.5 μL of the thrombin−hirudin complex and 1.5 μL of reservior solution containing: 0.05 M sodium phosphate buffer pH 7.3, 28% PEG8000. The soaking was carried out by adding powdered compound directly to the drop and leaving it at room temperature overnight.

Data Collection and Structure Determination. The diffraction data were collected from frozen crystals on a MarCCD detector mounted on a micro focus rotating anode generator FR-E Super-Bright from Rigaku. The crystal-to-detector distance was set to be 90 mm, and 360 images were collected with oscillation of 1°. The data were processed with MOSFLM[12] and programs in the CCP4 suite.[13] Refinement and model rebuilding was carried out using programs REFMAC5[14] and COOT.[15] The structure has been deposited with PDB, ID code 3DA9.

General Procedures. All chemicals were used as received, and solvents for dry conditions were kept over molecular sieves. All ^{1}H NMR spectra were recorded using a Varian Unity 500 at 500

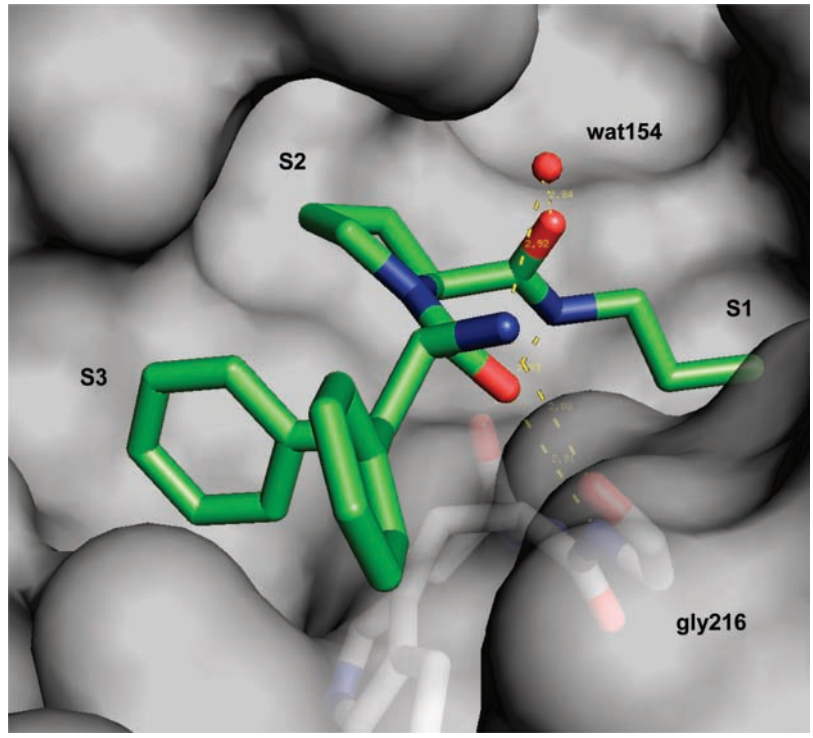


Figure 7. X-ray structure of compound **26** in the active site of thrombin with the H bonds to the enzyme backbone and wat154 highlighted (PDB, ID code 3DA9).

MHz. Flash column chromatography was carried out using Silica Gel 60 (0.063−0.200 mm). Semipreparative HPLC was performed using a Bishoff system, consisting of 2250 HPLC compact pumps and Lambda 1010 detector, equipped with a Sperisorb ODS2 (C18) 10 μ 250 mm × 10 mm or a Kromasil C18 10 μ 250 mm × 10 mm. Purification was performed either using different gradients of acetonitrile/H_2O, or isocratically. Normal phase semipreparative HPLC was performed on a Chromtech CT-sil 5 μ 250 mm × 10 mm isocratically using different proportions of heptane and isopropyl alcohol. Purity was checked on a HP 1090 using a Kromasil C18 5 μ 250 mm × 4.6 mm with a gradient ranging from 10 to 70% acetonitrile in 30 min, thereafter isocratically at 70% for an additional 15 min. Purity always exceeded 95% when monitoring the eluent with UV at 214 nm or 254 nm.

High resolution mass spectra was obtained on a Waters LCT Premier using positive electrospray for ion generation. Leucine Enkephaline $C_{28}H_{37}N_5O_7$ (*m*/*z* 556.2771) was used as reference.

Procedures and Data for 1−36. General procedures for A−D: 1 g of Boc-L-proline hydroxysuccinimide ester was dissolved in 30 mL dry THF. Three eqviv of the amine was added, and the mixture was stirred during 3 days at room temperature. The insoluble part was then filtered off, washed with THF, and the filtrate combined with the "wash solution" was concentrated in vacuo and subjected to flash chromatography using EtOAc as eluent. The product obtained was subsequently deprotected with freshly prepared 10 mL of AcCl/EtOH (1:2) for 1 h, evaporated into dryness, and left in a desiccator overnight. To afford the free base, the HCl salt was partitioned between 30 mL of 5 M NaOH and 3 × 30 mL of EtOAc or diethylether and dried over Na_2SO_4.

Pyrrolidine-2-carboxylic Acid Methylamide (A). Yield: 72%; ^{1}H NMR: 7.60 (s, 1H), 3.75 (dd, 1H, J = 7.5 Hz, J = 13.0 Hz), 3.02 (dd, 1H, J = 6.7 Hz, J = 15.5 Hz), 2.91 (m, 1H), 2.82 (d, 3H, J = 4.8 Hz), 2.15 (dt, 1H, J = 7.4 Hz, J = 15.3 Hz), 1.93 (td, 1H, J = 5.5 Hz, J = 12.4 Hz), 1.80 (br s, 1H), 1.72 (tt, 1H, J = 6.2 Hz, J = 12.3 Hz).

1-[3-(4-Chloro-phenyl)-propionyl]-pyrrolidine-2-carboxylic Acid Methylamide (1). As a general example: 46 mg of 3-(4-Chloro-phenyl)-propionic acid (1 equiv) were dissolved in 5 mL of ACN and 95 mg of HBTU (1 equiv) were added. After 3 min of activation, 100 μL (2 equiv) of DIPEA and 32 mg of A, dissolved in 5 mL ACN, were added. The reaction was allowed to proceed overnight. Work up was performed by partition between the ACN solution and 10 mL of saturated NaCl. The water solution was then extracted 3 × 10 mL EtOAc, the combined organic extracts was washed with 30 mL of 1 M HCl and 30 mL of saturated $NaHCO_3$ and was finally dried over Na_2SO_4. Purification performed on semipreparative HPLC using 60 H_2O/40 ACN on HPLC afforded **1** as clear oil in 78% yield. ^{1}H NMR: 7.25 (m, 2H), 7.16 (d, 2H, J = 8.3 Hz), 6.93 (s, 1H), 4.55 (d, 1H, J = 8.0 Hz), 3.45 (m, 1H), 3.31 (m, 1H), 2.96 (m, 2H), 2.76 (d, 3H, J = 4.8 Hz), 2.60 (m, 2H), 2.44 (m, 1H), 2.09 (m, 1H), 1.95 (m, 1H), 1.76 (m, 1H). HRMS found 295.1220; calcd 295.1213.

1-(3,3-Diphenyl-propionyl)-pyrrolidine-2-carboxylic Acid Methylamide (5). As a general example: 72 mg of 3,3-diphenyl-propionic acid (1 equiv) were dissolved in 20 mL of DMF and 121 mg of HBTU (1 equiv) were added. After 3 min of activation, 115 μL (2 equiv) of DIPEA and 41 mg of A, dissolved in 5 mL of DMF, were added. The reaction was allowed to proceed overnight. The mixture was subsequently concentrated in vacuo, and purification was performed on semipreparative HPLC using 65 H_2O/35 ACN on HPLC, affording white crystals in 70% yield. ^{1}H NMR: 7.25 (m, 10H), 6.55 (s, 1H), 4.68 (t, 1H, J = 7.8 Hz), 4.53 (d, 1H, J = 7.8 Hz), 3.44 (dt, 1H, J = 2.5 Hz, J = 9.4 Hz), 3.32 (dt, 1H, J = 7.3 Hz, J = 9.5 Hz), 3.06 (m, 2H), 2.60 (d, 3H, J = 4.8 Hz), 2.39 (m, 1H), 1.97 (m, 1H), 1.86 (m, 1H), 1.70 (tt, 1H, J = 7.3 Hz, J = 11.9 Hz). HRMS found 337.1932; calcd 337.1916.

1-[2-Amino-3-(4-chloro-phenyl)-propionyl]-pyrrolidine-2-carboxylic Acid Methylamide (21). First, 86 mg of D-2-Boc-3-(4-chloro-phenyl)-propionic acid were dissolved in 5 mL of ACN. Then 80 mg of HBTU were used to activate the acid for 3 min. Thereafter, 100 μL of DIPEA and 27 mg of A dissolved in 5 mL of ACN were added. The reaction was allowed to proceed overnight. Work up was performed by partition between the ACN solution and 10 mL of saturated NaCl. The water solution was then extracted with 3 × 10 mL EtOAc, the combined organic extracts were washed with 30 mL of 1 M HCl and 30 mL of saturated $NaHCO_3$, and were finally dried over Na_2SO_4. Purification was performed using gradient elution, starting with 70% H_2O to 33% H_2O in 34 min on semipreparative HPLC. The product obtained was subsequently deprotected with freshly prepared 2 mL of AcCl/EtOH (1:

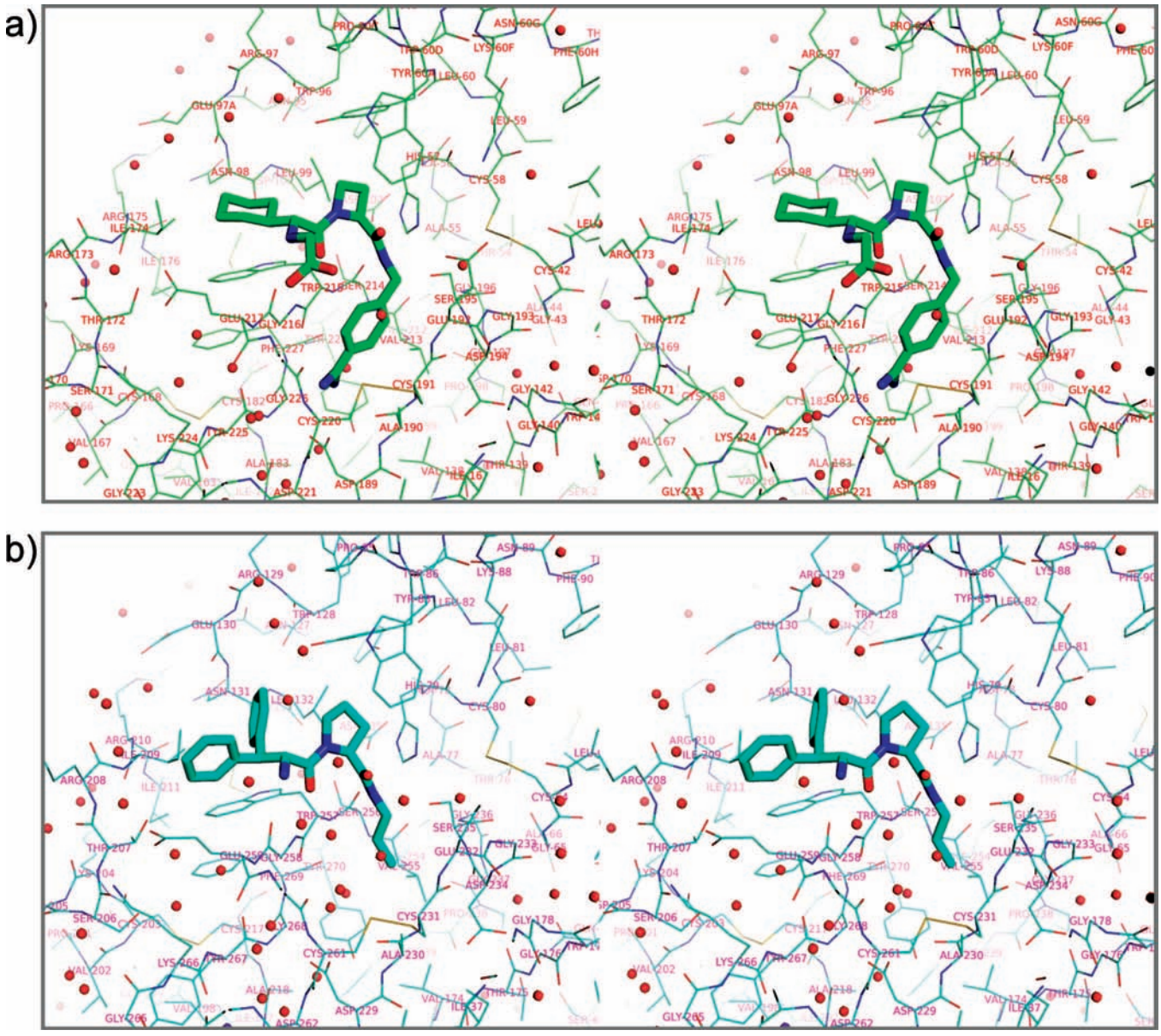


Figure 8. Stereo pictures of (a) melagatran (pdb:1Kk22) and (b) compound **26** (pdb: 3DA9) binding to active site of thrombin.

2) for 1 h, evaporated into dryness, and left in a desiccator overnight. To afford the free base, the HCl salt was partitioned between 5 mL 5 M NaOH and 3 × 10 mL EtOAc or diethylether and dried over Na_2SO_4. Yield 60% as transparent semi solid. ^{1}H NMR: 7.29 (d, 2H) 7.14 (d, 2H, J = 8.3 Hz) 6.95 (s, 1H) 4.46 (dd, 1H, J = 1.5 Hz, J = 8.2 Hz) 3.75 (t, 1H, J = 7.2 Hz) 3.51 (dt, 1H, J = 2.8 Hz, J = 9.0 Hz) 2.92 (m, 2H) 2.79 (m, 4H) 2.36 (m, 1H) 2.00 (m, 1H) 1.73 (m, 1H) 1.60 (s, 1H). HRMS found 310.1325; calcd 310.1322.

1-(2-Amino-2-cyclohexyl-acetyl)-pyrrolidine-2-carboxylic acid isobutyl-amide (33). Pyrrolidine-2-carboxylic acid isopropylamide was synthesized according to A using Boc-L-proline hydroxysuccinimde and isobutylamine. Synthesis and work up of the final product was performed according to **21** using the obtained pyrrolidine-2-carboxylic acid isopropylamide and Boc-D-cyclohexylglycine purified using 45 H_2O/55 ACN on HPLC, yielding 57% as transparent semisolid. ^{1}H-NMR: 7.20 (s, 1H), 4.59 (d, 1H, J = 7.5 Hz), 3.64 (dt, 1H, J = 2.7 Hz, J = 9.4 Hz), 3.49 (dt, 1H, J = 7.5 Hz, J = 9.5 Hz), 3.28 (d, 1H, J = 7.0 Hz), 3.05 (m, 2H), 2.46 (m, 1H), 2.12 (m, 1H), 1.96 (m, 2H), 1.78 (m, 5H), 1.52 (m, 2H), 1.26 (m, 2H), 1.10 (m, 2H), 0.99 (dq, 1H, J = 2.9 Hz, J = 12.3 Hz), 0.89 (d, 6H, J = 6.7 Hz). HRMS found 310.2500; calcd: 310.2495.

1-(2-Cyclohexyl-2-phenylmethanesulfonylamino-acetyl)-pyrrolidine-2-carboxylic Acid Methylamide (35). First, 34 mg of **29** were dissolved in 5 mL of DCM and cooled on ice. Thereafter, 3 equiv of triethylamine (56 μL) and 3 equiv of phenylmethanesulfonylchloride (77 mg) were added and the mixture was allowed to stir overnight. The mixture was concentrated in vacuo and purified using 60 H_2O/40 ACN on HPLC, yielding 69% as transparent semisolid. ^{1}H NMR: 7.45 (m, 2H), 7.39 (m, 2H), 6.81 (s, 1H), 4.90 (d, 1H, J = 9.7 Hz), 4.52 (dd, 1H, J = 1.9 Hz, J = 7.8 Hz), 4.29 (q, 2H, J = 13.9 Hz), 3.60 (dd, 1H, J = 7.9 Hz, J = 9.1 Hz), 3.49 (m, 1H), 3.31 (td, 1H, J = 7.9 Hz, J = 15.9 Hz), 2.78 (d, 3H, J = 4.8 Hz), 2.34 (m, 1H), 1.98 (m, 3H), 1.87 (d, 1H, J = 12.9 Hz), 1.72 (m, 3H), 1.52 (m, 2H), 1.13 (m, 4H), 0.90 (dq, 1H, J = 3.3 Hz, J = 12.5 Hz). HRMS found 422.2117; calculated 422.211.

Acknowledgment. We thank Maria Ericsson at AstraZeneca R&D Mölndal, Sweden, for providing us with results from the enzyme assay, Gary Franklin regarding improvement of the language, and finally, the University of Kalmar for funding.

Supporting Information Available: Synthesis and HRMS and H NMR identification of all other compounds. Chromatographic

analyses of compound purity. Additional information on X-ray structure determination. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Leung, D.; Abbenante, G.; Fairlie, D. P. Protease inhibitors: current status and future prospects. *J. Med. Chem.* **2000**, *43*, 305–341.

(2) Ilies, M. A.; Scozzafava, A.; Supuran, C. T. Therapeutic applications of serine protease inhibitors. *Expert Opin. Ther. Pat.* **2002**, *12*, 1181–1214.

(3) Maryanoff, B. E. Inhibitors of serine proteases as potential therapeutic agents: the road from thrombin to tryptase to cathepsin G. *J. Med. Chem.* **2004**, *47*, 1–19.

(4) Abbenante, G.; Fairlie, D. P. Protease inhibitors in the clinic. *Med. Chem.* **2005**, *1*, 71–104.

(5) Sorbera, L. A.; Bayes, M. C. J.; Silvestre, J. Melagatran and Ximelagatran. *Drugs Future* **2001**, *26*, 1155–1170.

(6) Lumma, W. C., Jr.; Witherup, K. M.; Tucker, T. J.; Brady, S. F.; Sisko, J. T.; Naylor- Olsen, A. M.; Lewis, S. D.; Lucas, B. J.; Vacca, J. P. Design of novel, potent, noncovalent inhibitors of thrombin with nonbasic P-1 substructures: Rapid structure−activity studies by solid phase synthesis. *J. Med. Chem.* **1998**, *41*, 1011–1013.

(7) Huber, W.; Mueller, F. Biomolecular interaction analysis in drug discovery using surface plasmon resonance technology. *Curr. Pharm. Des.* **2006**, *23*, 3999–4021.

(8) Jagabandhu, D.; Kimball, S. D.; Hall, S. E.; Han, W.-C.; Iwanowics, E.; Lin, J.; Moquin, R. V.; Reid, J. A.; Sack, J. S.; Malley, M. F.; Chang, C. Y.; Chong, S.; Wang-Iverson, D. B.; Roberts, D. G. M.; Seiler, S. M.; Schumacher, W. A.; Ogletree, M. L. Molecular Design and Structure−Activity Relationships Leading to the Potent, Selective, and Orally Active Thrombin Active Site Inhibitor BMS-189664. *Biorg. Med. Chem. Lett.* **2002**, *12*, 45–49.

(9) Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discovery* **2007**, *6*, 211–219.

(10) Trygg, J.; Wold, S. Orthogonal projections to latent structures (OPLS). *J. Chemom.* **2002**, *16*, 283–293.

(11) Nilsson, J. W.; Kvarnstrom, I. D.; Musil, D.; Nilsson, I.; Samulesson, B. Synthesis and SAR of thrombin inhibitors incorporating a novel 4-amino-morpholinone scaffold: analysis of X-ray crystal structure of enzyme−inhibitor complex. *J. Med. Chem.* **2003**, *46*, 3985.

(12) Leslie, A. G. Integration of macromolecular diffraction data. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 1696–1702.

(13) Collaborative Computational Project no. 4. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *50*, 760−763.

(14) Murshudov, G. N.; Vagin, A. A.; Lebedev, A.; Wilson, K. S.; Dodson, E. J. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 247–255.

(15) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.

JM8011849