

## Articles

### Generation of Ligand Conformations in Continuum Solvent Consistent with Protein Active Site Topology: Application to Thrombin

Paulette A. Greenidge,<sup>‡</sup> Sandrine A. M. Mérette,<sup>‡</sup> Richard Beck,<sup>‡</sup> Guy Dodson,<sup>‡</sup> Christopher A. Goodwin,<sup>#</sup> Michael F. Scully,<sup>#</sup> John Spencer,<sup>‡</sup> Jörg Weiser,<sup>§</sup> and John J. Deadman<sup>\*,‡</sup>

Chemistry Department, Drug Discovery Division, and Biochemistry Department, Thrombosis Research Institute, Emanuelle Kaye Building, Manresa Road, London, SW3 6LR. U.K., Anterio Consult & Research GmbH, Augustaanlage 26, 68165 Mannheim, Germany, and Protein Structure Division, National Institute for Medical Research, The Ridgeway, Mill Hill, London, NW7 1AA. U.K.

Received August 20, 2002

Using the crystal structure of an inhibitor complexed with the serine protease thrombin (PDB code 1UVT) and the functional group definitions contained within the Catalyst software, a representation of the enzyme's active site was produced (structure-based pharmacophore model). A training set of 16 homologous non-peptide inhibitors whose conformations had been generated in continuum solvent (MacroModel) and clustered into conformational families (XCluster) was regressed against this pharmacophore so as to obtain a 3D-QSAR model. To test the robustness of the resulting QSAR model, the synthesis of a series of non-peptide thrombin inhibitors based on arylsophonyl derivatives of an aminophenol ring linked to a pyridyl-based S1 binding group was undertaken. These compounds served as a test set (**20–24**). The crystal structure for the novel symmetrical disulfonyl compound **24**, in complex with thrombin, has been solved. Its calculated binding mode is in general agreement with the crystallographically observed one, and the predicted  $K_i$  value is in close accord with the experimental value.

#### Introduction

Thrombin (EC 3.4.21.5) has been implicated in the aetiology of a number of disease states including thrombotic disease,<sup>1</sup> cancer, and endotoxic shock, and for this reason inhibition of its proteolytic activity continues to be the subject of inhibitor design after three decades.<sup>2</sup> Second generation, nonpeptidic thrombin inhibitors are now in phase II<sup>3</sup> and phase III<sup>4</sup> clinical development.

Pooling the combined research in the thrombin area provides numerous libraries of inhibitors with diverse structures and a range of affinities for thrombin spanning 9 orders of magnitude (millimolar to picomolar). In addition, there are at least 124 structures of inhibitor complexes with thrombin deposited in the Protein Data Bank (Merops database, www.MEROPS.Sanger.Ac.Uk, code S01.217). Thrombin, in contrast to the other coagulation proteinases, shows only small changes in catalytic constants for its substrates upon binding to cofactors. For example, binding of thrombin to glycoprotein 1b<sup>5</sup> increases  $k_{cat}/K_m$  for cleavage of PAR-1 by only 6-fold,<sup>6</sup> while binding of FXa to FVa or FIXa to FVIIIa cause an increase of 3 orders of magnitude.<sup>7</sup> This small scale of allosteric change for thrombin is also reflected in its interactions with thrombomodulin, fi-

brinogen, the hirudin peptide, and sodium binding<sup>8</sup> and in the interactions of its mutants with these ligands and reflects only small movements of the protein ternary structure.<sup>9</sup> For this reason, the structures with diverse ligands should be comparable, and indeed it has even been possible to identify conserved waters.<sup>10</sup> Several approaches<sup>11,12</sup> have been described to rationalize this extensive data and to allow extrapolation to the design of new inhibitors. For example, the comparative molecular similarity indices approach (CoMSIA)<sup>11</sup> was applied, but this method does not consider the structure of the receptor. In contrast, the "linear response"<sup>12</sup> method which does utilize the target was used to develop several equations to predict thrombin inhibitor affinities. While this method allows a range of factors that are important to inhibitor–enzyme interactions to be included, the Monte Carlo method is inherently computationally expensive.

Bursi and Grootenhuys<sup>13</sup> have previously correlated theoretical and experimentally determined binding data for a series of thrombin inhibitors. They found that molecular mechanics minimization of inhibitors in the receptor structure in which they had been cocrystallized only provided a statistically significant correlation ( $R = 0.74$ , 14 inhibitors) for those complexes with a resolution of  $\leq 2.5$  Å. Comparative molecular field analysis (CoMFA)<sup>14</sup> which only uses the information of the inhibitors performed better ( $R = 0.95$ , cross-validated  $r^2$  ( $q^2$ ) = 0.46) when used with high quality ab initio charges and minimized conformations of crystal structure inhibitors (crystal structure alignment). Thus, in this case, the success of both the aforementioned

\* To whom correspondence should be addressed. Tel: ++44+207-351 8330 ext. 3622. Fax: ++44+207-351 8324. Email: jdeadman@trigen.co.uk.

<sup>‡</sup> Chemistry Department, Drug Discovery Division, Thrombosis Research Institute.

<sup>#</sup> Biochemistry Department, Thrombosis Research Institute.

<sup>§</sup> Anterio Consult & Research GmbH.

<sup>‡</sup> National Institute for Medical Research.

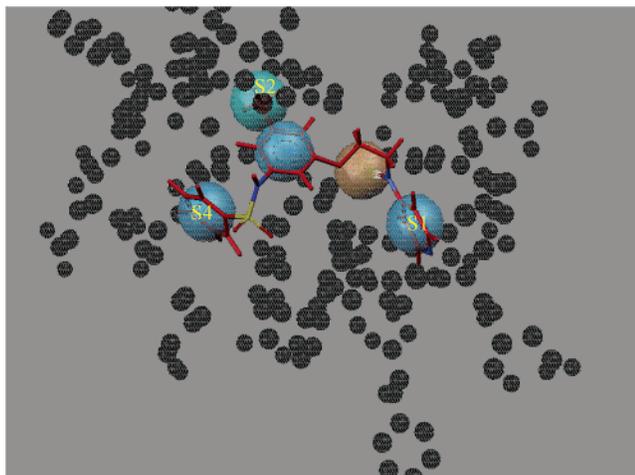
**Table 1.** Structure, Experimental (exptl), Calculated (calcd)  $K_i$  Values, and Catalyst Conformational Energies of the 16 Training Set Inhibitors Plus a Test Set Inhibitor Taken from the Literature (Compound 17)

No	Structure	$K_{iexptl}$ ( $\mu$ M)	$K_{icalcd}$ ( $\mu$ M)	Conformational Energy(kJ/mol)	No	Structure	$K_{iexptl}$ ( $\mu$ M)	$K_{icalcd}$ ( $\mu$ M)	Conformational Energy(kJ/mol)
2		0.004	0.020	14.5	9		0.316	0.29	0.2
3		0.005	0.025	29.0	10		0.316	0.68	6.2
4		0.006	0.013	34.8	11		0.316	0.70	5.7
1		0.025	0.015	6.6	12		0.631	0.18	0.4
5		0.04	0.082	8.5	13		0.631	0.86	9.2
6		0.05	0.036	0.2	14		1.585	0.60	10.0
7		0.063	0.80	55.6	15		1.585	0.26	7.9
8		0.158	0.097	0.7	16		1.995	0.053	8.5
					17		4	0.95	<84

methods apparently relies on access to crystal structure data for each inhibitor–enzyme complex. However, this dependence makes these approaches impracticable during the early stages of a medicinal chemistry program; hence, our investigation of the novel combination of a structure-based pharmacophore<sup>15</sup> (representation of important inhibitor–enzyme interactions in which excluded volumes that are not penetrable by the inhibitors are used to define the demarcation of the active site) constructed from a single high-resolution crystal structure complex, the Catalyst<sup>16</sup> pharmacophore mapping method, and in continuum solvent conformations of ligands generated in MacroModel.<sup>17</sup> The work described here considers a subset of that same set of thrombin inhibitors.<sup>13</sup> The structure-based pharmacophore (Figure 1) was built using the crystallographic coordinates of compound **1** (Table 1) complexed with thrombin (PDB code 1UVT)<sup>18</sup> to predict the effects on the  $K_i$  values of structural modification of a set of homologous 4-aminopyridine (4-AP) thrombin inhibitors (Table 1).

In Catalyst,<sup>16</sup> a conformational model is an abstract representation of the accessible conformational space of a ligand. It is assumed that the biologically active conformation of a ligand (or a close approximation thereof) should be contained within this model. Given a ligand training set (ligand conformational models and binding affinity values) and pharmacophore model,

Catalyst<sup>16</sup> attempts to produce a 3D-QSAR model that correlates estimated activities with measured activities. A pharmacophore model consists of a collection of features necessary for the biological activity of the training set arranged in 3D-space. Common pharmacophore features include hydrogen bond donor, hydrogen bond acceptor, and hydrophobe. The fewer features an inhibitor maps to, and the poorer its fit to them, then the lower its affinity will be predicted to be. The simplistic scoring function does not explicitly contain terms for ligand solvation nor protein and ligand entropy; thus its use may be limited to congeneric ligand series.<sup>19</sup> The work of Vieth et al.<sup>20</sup> supports the pharmacophore approach<sup>21</sup> and the use of conformational models generated in solution. By means of molecular dynamics simulations, atoms that play a key role in ligands binding affinity were identified as those having the lowest mobility and largest interaction energy with the receptor. These so-called “anchor” atoms have a similar spatial orientation both in the active site and in solution. In an analysis of conformational energy penalties of protein-bound ligands, Boström et al.<sup>22</sup> also concluded that it is more appropriate to use the aqueous rather than the in vacuo conformational ensemble of the unbound ligand when making a comparison with the bioactive conformation. Several recent studies,<sup>23</sup> including our own,<sup>24</sup> have appeared in which the conforma-



**Figure 1.** Catalyst mapping of the crystal structure conformation of compound **1** with thrombin to the structure-based pharmacophore. Three aromatic-hydrophobic features (dark blue spheres) correspond to the three aromatic rings of compound **1** and a general hydrophobic feature (light blue) to its methyl group. A hydrogen-bond donating feature (cyan) is derived from the interaction of the 4-aminopyridine group with the carbonyl oxygen atom of amino acid residue Ser-214 (chymotrypsin numbering scheme). The excluded (ligand-inaccessible) volumes (black spheres) correspond to the atoms delimiting the active site of thrombin. The important regions of the enzyme active site, the specificity pocket of thrombin, S1, and the two hydrophobic pockets S2 and S4 are labeled.

tions of molecules in aqueous solution have been used to establish a pharmacophore model. In our previous study,<sup>24</sup> conformations were generated in continuum solvent (MacroModel)<sup>17</sup> for this data set (Table 1) in which ligands contain up to 14 rotatable bonds. Each conformational model typically consisted of several thousand conformers, thus making the calculation of the mapping mode to the pharmacophore model slow. However, by clustering the results of a conformational search into conformational families with the program XCluster Version 1.7,<sup>25</sup> the number of conformers has now been reduced to a more tractable number. Hence, in continuum solvent generated conformational models are now computationally amenable to use for pharmacophore mapping of large data sets. A crystal structure for a newly synthesized inhibitor cocrystallized with thrombin is presented and it is checked whether the highest scoring pharmacophore binding mode is indeed consistent with available crystal structure data.

## Methods

**Catalyst Pharmacophore Construction and 3D-QSAR. Structure-Based Pharmacophores.** The location of features in the structure-based pharmacophores were defined by the crystallographic coordinates of atoms in the BM14.1248/thrombin complex (PDB reference code 1UVT) in which water molecules had been removed.

### Interactions of Compound **1** with Thrombin.

(i) The 4-amino group is hydrogen bonded directly to the Ser-214 (chymotrypsin numbering) carbonyl oxygen atom.

(ii) The proton of the NH group of the pyridine ring forms a hydrogen bond with a carboxylate oxygen atom of Asp-189 in the S1-pocket.

(iii) The methyl group occupies the S2-pocket.

(iv) In the aryl binding site, the phenyl group of the inhibitor takes part in an edge-to-face aromatic stacking interaction with Trp-215 and is flanked by a number of other hydrophobic residues (S4-pocket).

The hydrogen bond of the pyridine ring NH group with Asp-189 was not included in the pharmacophore definition, but the centroid of the pyridine ring (common to the inhibitors studied) defined as a hydrophobic (S1-pocket) aromatic center instead since the major interactions in the S1-pocket are not simply driven by electrostatics. Noncovalent thrombin inhibitors that have nonbasic groups in the P1 position and that exhibit high affinity exist,<sup>26</sup> and it has been demonstrated that the 4-aminopyridine moiety contributes only negligibly to the total binding energy of such ligands to thrombin.<sup>27</sup> Consequently, all molecules were constructed with an unprotonated pyridine moiety. The centroids of the two remaining aromatic rings of the inhibitor were likewise used to define hydrophobic aromatic features. Thus, such a pharmacophore definition allows for alternative binding modes by not biasing the pyridine ring to occupy the S1-pocket. A precedent for unexpected "inverse" binding modes with peptidic thrombin inhibitors is known, where the interaction of an N-terminal trifluoroacetyl group with the enzyme catalytic triad was predominant.<sup>28</sup> The methyl carbon atom was defined as a hydrophobic feature (S2-pocket), and the Ser-214 carbonyl oxygen atom as a hydrogen bond donor feature. The other end of the vector was removed. The spherical tolerance for all features was 1.5 Å. BM14.1248 (compound **1**) was extracted from the enzyme/inhibitor complex, and this crystallographic conformation was stored. BM14.1248 was able to successfully map to all features of the described pharmacophore. Subsequently, the remaining atoms delimiting the active site were represented as excluded volumes (space which the inhibitor is not allowed to occupy) defined within a cut off of 6 Å from any atom of the inhibitor using Insight II.<sup>29</sup> One pharmacophore model was constructed in which the excluded volumes were scaled at 30% of their respective atomic van der Waals radii.<sup>30</sup>

**Training and Test Sets.** The training set (inhibitor conformational models and  $K_i$  values) for the regression of the structure-based pharmacophores consisted of 16 inhibitors taken from the data set of Bursi and Grootenhuis<sup>13</sup> (inhibitors with undefined stereochemistry were excluded) (Table 1). Sixteen compounds represent the minimum recommended number of molecules to be included in a training set. Compound **17** (Table 1) was selected as a test set inhibitor to examine the versatility of the structure-based pharmacophore (see Results and Discussion). The inhibitor/thrombin crystal structure complex is available for compound **17** (PDB reference code 1UVS). Note that care should be taken when importing conformational models generated external to Catalyst as the recalculation of conformational energies may result in some conformers being given a conformational energy above the default 84 kJ/mol threshold such that these conformers are excluded from the mapping process. However, all conformers of the training set had energies calculated by Catalyst to be less than 84 kJ/mol above the lowest energy conformer for a given compound. Three conformers of compound **17** were

considered to have conformational energies above the threshold. These were stored as separate molecules by using the "show conformational model" option, followed by "add to view". The conformer was then saved and the necessary conformational model was made by registering this conformer. The conformational model giving the best fit score for compound **17** is discussed in Results and Discussion. The catScramble script was used to randomize (five trials) the experimental values of the training set spreadsheet for regression with the pharmacophore. All calculations were performed using an Iris Indigo Elan, R4000, memory size 160 Mbytes, 100 MHz IP20 processor.

**MacroModel/XCluster.** The following protocol was used to generate conformations for each compound:

(1) Execute an exhaustive conformational search using the molecular modeling program MacroModel<sup>17</sup> (version 7.2). For calculations on all compounds the latest development of the current best general-purpose force-field for medicinal chemistry, MMFF94s,<sup>31</sup> was used in combination with the GB/SA solvation model.<sup>32</sup> GB/SA treats solvent as analytical dielectric continuum that starts near the van der Waals surface of the solute and extends to infinity. The model includes both generalized Born-based (GB) solvent polarization terms and surface area-based (SA)<sup>33</sup> solvent displacement terms. All nonbonded cutoffs were set to infinity for all calculations. Energy minimizations were performed with the Truncated Newton–Raphson conjugate gradient (TN-CG)<sup>34</sup> method, which involves the use of second derivatives; the derivative convergence criterion was set to 0.05 kJ/Å-mol. Conformational search was performed by the Monte Carlo<sup>35</sup> method for the random variation of all of the rotatable bonds combined with the so-called low mode conformational search (LMCS)<sup>36</sup> algorithm. For each calculation 10 000 Monte Carlo steps were carried out.

(2) Sort all found conformations according to energy.

(3) Store only conformations whose molecular mechanics energy differences to the calculated global energy minimum of a compound (GEM) are below 15 kJ/mol. The "exhaustive" conformational models (both in vacuo and in continuum solvent) in our earlier study were based on conformational energy thresholds of 42 kJ/mol. It was decided to decrease that energy window to avoid rather unpopulated conformations (according to the Boltzmann distribution) and distorted conformations being classified as conformational families, since the number of conformational families is restricted (see point 4 in this protocol). Additionally, this threshold is in line with the results of a study by Boström et al.<sup>22</sup> They found that for 70% of the protein–ligand complexes that they investigated, the energies of the bioactive conformations were within 12.5 kJ/mol of their respective minimum energy conformation. Uncertainties in the interpretation of the experimental data or limitations of the computational methods accounted for the higher calculated penalties in the remaining 30% of complexes.

(4) Cluster the remaining conformations in exactly 200 different conformational families (numbered from 1 to 200) based on atomic RMSD after rigid-body superposition using XCluster.<sup>25</sup> The conformational family with the number 1 always contains the GEM;

the numbering of all other conformational families is arbitrary with regard to conformational energies. By means of comparison with "exhaustive" conformational models (generated by a modified systematic search method in which torsion angles are varied in specified increments) and assessment with two quantitative metrics, it has been shown that a small number of conformations are sufficient to represent the low-energy conformational spaces of small- to medium-sized molecules (using the method of conformation generation as implemented in Catalyst).<sup>37</sup> This led to the derivation of a relationship between the number of rotatable bonds possessed by a molecule, the number of conformers and the resolution of a conformational model. The resolution of a conformational model (subset) can be defined in terms of the maximum expected rms distance of an arbitrary low-energy conformer extracted from an exhaustive set (larger set), from a conformation in the model.<sup>37</sup> The number of rotatable bonds possessed by molecules in this data set is in the main between 7 and 10, the exceptions being compound **2** with 14 (and compound **17** with 12). Accordingly, a Catalyst generated conformational model size of 200 is appropriate for the choice of tolerance sphere radius (1.5 Å). For this reason, we also chose a MacroModel cluster size of 200.

(5) Store the leading conformation (conformation with the lowest energy) of each conformational family. According to the previous point of this protocol the leading conformation of the first conformational family is always the GEM.

(6) For each compound write one multiple coordinate file, which contains the 200 leading conformations in ascending order of the numbering of the conformational families. Conveniently the first conformation of such a multiple coordinate file is therefore always the GEM (see points 4 and 5 of this protocol). Each multiple coordinate file represents the clustered in continuum solvent conformational model of each compound. All calculations were performed either with the Linux version of MacroModel 7.2<sup>17</sup> on a 750 MHz Pentium III processor or the Sun version on a Sun Ultra 2 machine.

## Results and Discussion

**Comparison of "Exhaustive" and "Clustered" In Continuum Solvent Conformational Models: Impact on 3D-QSAR Model.** The crystal bound conformation of compound **1** (BM14.1248) with thrombin (PDB code 1UVT), and its interactions therewith, were used to define a structure-based pharmacophore (Methods, Figure 1). The excluded volumes (space which the inhibitor is not allowed to penetrate and defined by atoms delimiting the active site) were scaled to 30% of their respective van der Waals radii to allow for direct comparison with an earlier study.<sup>24</sup> Scaling may be a necessary consequence of the assumption of a rigid enzyme structure as Murray et al.<sup>38</sup> have shown that for a given enzyme (the data set included thrombin) the docking success of flexible ligands is reduced against enzyme crystal structures derived from other ligands. In that earlier study,<sup>24</sup> "exhaustive" conformational models for thrombin inhibitors (Table 1) both in vacuo and in continuum solvent with the MacroModel<sup>17</sup> software were produced, as well as a conformational model using Catalyst ("best" mode). Three criteria were uti-

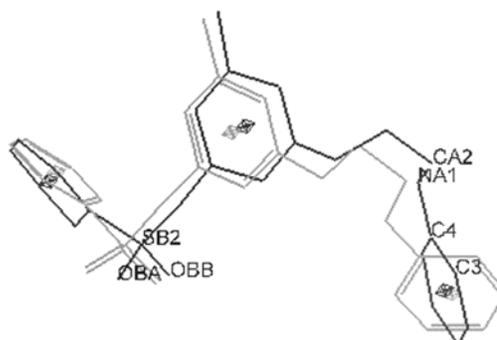
lized to assess the quality of the resulting 3D-QSAR models following regression of the structure-based pharmacophore (Figure 1) with these collections of conformers. On the basis of the correlation coefficient and root-mean-square deviation (RMSD) between estimated and observed  $K_i$  values, and the ability to reproduce crystallographically observed binding modes, the continuum solvent conformational model was judged to produce the best 3D-QSAR model with the Catalyst generated conformers performing least well.<sup>24</sup> For instance, compound **1** was calculated to bind in an inverse binding mode. The S4-feature was not mapped and the benzene-sulfone moiety occupied the S1 pocket instead of the pyridine ring as observed in the crystal structure complex (1UVT). However, the closely related compound **9** assumed the expected binding mode and mapped four of the five pharmacophore features. Given the conservative nature of the training set the inconsistency of these results was disappointing. Both the Macromodel<sup>17</sup> "exhaustive" conformational models for compound **1** resulted in a mapping mode such that all five features of the pharmacophore were satisfied as in the crystal structure complex (1UVT). Each "exhaustive" conformational model consisted of up to several thousand conformers, making them slow to import into Catalyst and in performing calculations thereafter; however, the size of the conformational models has now been reduced by clustering them into conformational families (see Methods).

In the current work, after regression of the pharmacophore model (Figure 1) with the training set ("clustered" in continuum solvent conformational model, Table 1) the correlation coefficient and root-mean-square deviation (RMSD) between estimated and observed  $K_i$  values have been examined. As compared to the 3D-QSAR model regressed with the "exhaustive" in continuum solvent conformational model the quality of the present model is degraded with  $R$  values of 0.69 and 0.75 and RMSD values of 1.32 and 1.21, respectively. Interestingly, the figures for the clustered in continuum solvent conformational model still compare favorably with those for the "exhaustive" in vacuo conformational model for which  $R = 0.57$  and  $RMSD = 1.52$ .<sup>24</sup> See point (vii) below for a discussion as to why this is so. In the present study, compound **1** does not map to the hydrogen bond feature defined by Ser-214 (Table 2), although it does map to all four of the hydrophobic features. Figure 2 shows the superimposition of the crystal structure of compound **1** extracted from the complex with thrombin and the conformation that maps to the pharmacophore with "fast" fit. A RMSD value of 1.97 Å is obtained over 36 atoms using the centroids of the aromatic rings for superimposition. The greatest conformational variation between the conformers occurs about the torsion defined by the atoms CA2 NA1 C4 C3 (Figure 2). A value of 178.4° is obtained for the crystal structure conformation and this compares with an equivalent value of 2.74° for the conformation in continuum solvent. Hence, the inability of the generated conformation of compound **1** to map to the hydrogen bond donor feature as the 4-amino hydrogen atom points away from the Ser-214 carbonyl oxygen atom. Although a conformer of compound **1** that is able to map to all five features of the pharmacophore model does exist, in

**Table 2.** A Comparison of the Structure-Based Pharmacophore (Figure 1) Features<sup>a</sup> Mapped by the Training Set Inhibitors Whose Conformational Models Are Generated In Continuum Solvent

compound	S1	HBD	S2	central phenyl ring	S4
<b>1</b>	+	-	+	+	+
<b>2</b>	+	-	+	+	+
<b>3</b>	+	-	+	+	+
<b>4</b>	+	-	+	+	+
<b>5</b>	+	-	+	+	+
<b>6</b>	+	+	+	+	-
<b>7</b>	+	+	-	+	-
<b>8</b>	+	+	+	+	-
<b>9</b>	+	+	-	+	+
<b>10</b>	+	+	-	+	-
<b>11</b>	+	+	-	-	+
<b>12</b>	+	+	-	+	+
<b>13</b>	+	-	-	+	+
<b>14</b>	+	-	-	+	-
<b>15</b>	+	+	-	+	+
<b>16</b>	+	+	-	+	+

<sup>a</sup> S1, S2, and S4 refer to hydrophobic features corresponding to the S1-, S2-, and S4-pockets respectively of thrombin. HBD refers to the hydrogen bond donor feature defined by the Ser-214 carbonyl oxygen atom and the centroid of the central phenyl ring of compound **1** (Table 1) is used to define a hydrophobic feature. Note that no indication is given of how well a particular feature is mapped by an inhibitor.



**Figure 2.** Superimposition of the crystal structure conformation of compound **1** (bold) extracted from the complex with thrombin (1UVT) and the conformation calculated to have the best fit score to the pharmacophore model (light shade).

doing so it does not obtain the best fit score. This finding may perhaps be attributed to two factors. First, that the current Catalyst fit score may not be a perfect measure of conformational quality, and second the choice of tolerance radii. It may well be the case that the tolerance radii for the pharmacophore regressed with the "clustered" conformational model should be made larger as compared to the pharmacophore regressed with the "exhaustive" conformational models.

**Structure-Based Pharmacophore Model and 3D-QSAR.** Catalyst assumes that all chemical features contribute equally in providing binding energy, and that compounds are more or less active because they possess or do not possess features that contribute positively to activity. By use of Tables 1 and 2 the mapping mode of the ligands and the consequent effect on calculated  $K_i$  values can be assessed. The contribution of the different moieties of which compounds such as **1** are comprised, to binding energy with thrombin has been examined. The inhibition constant increases considerably when the arylsulfonamide moiety is removed, and compounds that lack the central phenyl group do not inhibit thrombin at all. As compared to benzamidine the pyridine ring

**Table 3.** Randomization of the Experimental  $K_i$  Data to Test the Robustness of the 3D-QSAR Model

data	$R^a$	RMSD
nonscrambled	0.69	1.32
scrambled		
trial 1	-0.12	1.94
trial 2	-0.04	1.93
trial 3	0.50	1.60
trial 4	0.1	1.90
trial 5	-0.1	1.94

<sup>a</sup> Experimental  $K_i$  values ( $K_{i\text{expt}}$ , Table 1) were randomized using catScramble (see Methods) and the table shows the results of correlation with calculated  $K_i$  Values in five trials.

penetrates deeper into the S1-pocket, but is a less significant source of binding energy.<sup>27</sup>

Given that the pharmacophore features are based on the interactions of compound **1** with thrombin, it is not surprising that the most potent compounds, **2**, **3**, and **4** are calculated to have affinities less than or approximately equal to that of compound **1** even though they have the potential to make additional interactions with thrombin (Table 2). Thus, an obvious shortcoming of the Catalyst 3D-QSAR approach is that a predefined pharmacophore description is required. To generalize the pharmacophore, more features could be included to represent interactions thought to be important for inhibitors binding to thrombin,<sup>39</sup> but not necessarily made by this particular series of compounds. For example, in a series analogous to the current one, but which has a P1 guanidinoalkyl moiety, the sulfonamide NH group forms a weak hydrogen bond to the Gly-216 carbonyl oxygen atom.<sup>40</sup> However, the sulfonamide moiety of the present series is not assumed to directly interact with the protein.

The pharmacophore model does seem to reflect the trends of compound **9** versus:

(i) Compound **1**. Compound **9** lacks a group capable of mapping to the hydrophobe corresponding to the S2-pocket (Table 2).

(ii) Compound **5**, which has as an R2-substituent as a chlorine atom (Table 1), and in which the sulfonamide nitrogen atom is methylated.

(iii) Compounds **6**, **8**, and **10**. Given that the pharmacophore is defined using compound **1** which has a sulfonamide linker as compared to the shorter sulfone linker common to these three compounds (Table 1), it is not surprising that they are unable to map to the S4-feature (Table 2).

(iv) Compound **7** which is in good agreement with our measured  $K_i$  value (Table 3).

(v) Compound **11**, which has the same  $K_i$  observed as compound **9** (Table 1).

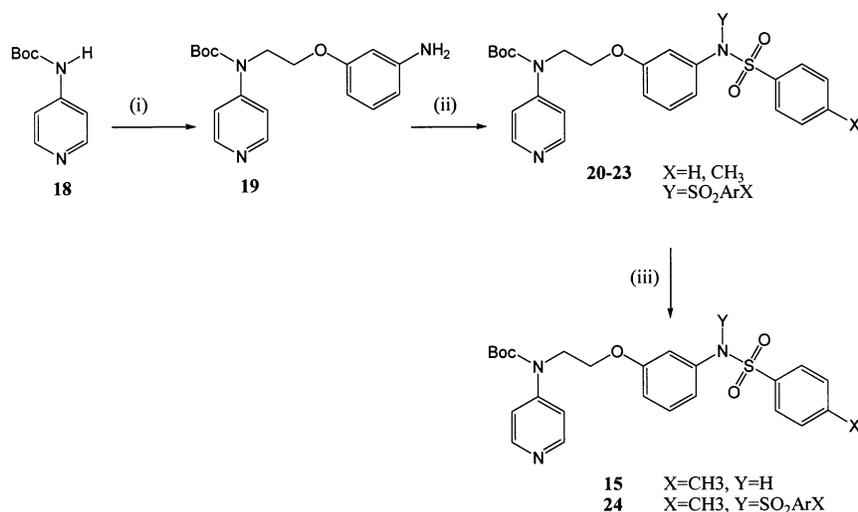
(vi) Compounds **12** and **15** for which excluded volume problems may play a role (increasing the size of the excluded volumes may increase the sensitivity of the model).

(vii) Compound **13**, which is unable to make a hydrogen bond with Ser-214 as both its sulfonamide and 4-amino nitrogen atoms are methylated (Tables 1 and 2), and compound **14** in which the sulfonamide group is reversed relative to compound **9**. As compared to compound **9** the sulfonamide linker of compound **14** is reversed (Table 1), and this is conformationally significant as the sulfone group is coplanar to the aromatic ring to which it is directly attached. Whereas in

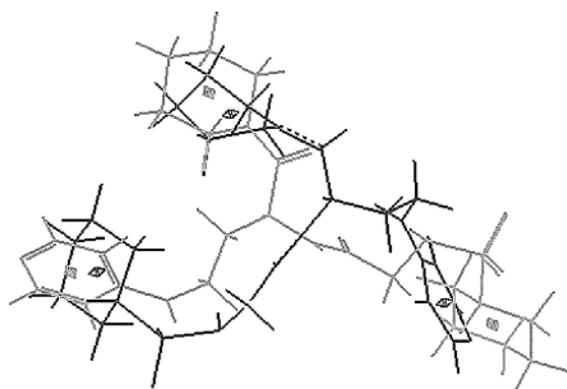
continuum solvent the minimum energy conformer of compound **9** is relatively extended, the equivalent conformer of compound **14** is folded such that the benzene-sulfonamide ring and central aromatic ring engage in a face to face pi-stacking interaction. To give an indication of the difference in conformation between these two molecules, the distance between the pyridine ring nitrogen atom and para-carbon atom of the benzene-sulfonamide moiety is given. The values for compounds **9** and **14** are 15.6 and 10.5 Å, respectively. So what in two dimensions appears to the eye to be a minor structural difference between compounds **9** and **14** (Table 1) in fact has a major impact on their  $K_i$  values (Table 1), perhaps explained by the inability of compound **14** to occupy the S4-pocket (Table 2). Using a 42 kJ/mol energy threshold, a conformer of compound **14** exists in the in vacuo conformational model such that the affinity of compound **14** is overestimated ( $K_{i\text{expt}}/K_{i\text{calcd}} = 43$ ). This contributes to the lower  $R$  and higher RMSD values of the resulting 3D-QSAR model as compared to that resulting from regression of either of the continuum solvent conformational models (exhaustive or clustered) with the pharmacophore model.

However, the model fares less well with respect to compound **16** which has a benzylamine group instead of a sulfonamide group as possessed by a majority of the other molecules. A possible enhancement to the pharmacophore might include describing the S4-pocket as an aromatic feature instead of as a hydrophobe, thus creating a more realistic description of the geometric relationship between the phenyl and Trp-215 rings (see Method point (iv)).

To test the robustness of the 3D-QSAR model, the pharmacophore was regressed several times with scrambled activities (Table 3). The correlation coefficients for the randomized experimental data were less than for the nonscrambled data and the RMSD values higher. Thus, the best 3D-QSAR model is obtained from correlating the calculated  $K_i$  values for the compounds with their corresponding experimental  $K_i$  values. In addition, before deciding if the model should prove useful in guiding further synthesis a check should be made whether compounds external to the training set map to the pharmacophore in chemically reasonable ways<sup>41</sup> as it has been argued that validation of a QSAR model can only be accomplished by use of an external set of compounds.<sup>42</sup> For this reason, the literature compound **17** (Table 2) which although related to compound **1** does show some structural variance was chosen as part of the test set. In the crystal structure with thrombin, compound **17** adopts an "Argatroban-like" conformation such that the N-carboxypiperidyl (S2-pocket) and cyclohexyl sulfonyl (S4-site) rings are in close proximity and thus able to make a number of van der Waals interactions with each other.<sup>43</sup> The pyridine ring occupies the S1-site, but the 4-amino group is unable to form a hydrogen bond with the Ser-214 amino acid residue. However, the mapping of the continuum solvent generated conformation of compound **17** to the pharmacophore predicts an "inverse" binding mode relative to that observed in the crystallographic complex with thrombin. The pyridine ring maps to the S4-hydrophobe, the N-carboxypiperidyl ring occupies a region close to the S1-hydrophobe, allowing the sulfona-

Scheme 1<sup>a</sup>

<sup>a</sup> Reagents and conditions: (i) (a) ethylene glycol di-*p*-tosylate, DMF, NaH, (b) 3-aminophenol, NaH; (ii) benzenesulfonyl chloride TEA, DCM; (iii) 50% TFA in DCM.



**Figure 3.** Superimposition of the crystal structure conformation of compound **17** (bold) extracted from the complex with thrombin (1UVS) and the conformation calculated to have the best fit score to the pharmacophore model (light shade).

amide group to map to the Ser-214 hydrogen bond donor feature and the N-carboxy piperidyl to the S2-pocket. The RMSD between the highest scoring conformer of compound **17** that maps to the pharmacophore model and the crystallographic conformation in complex with thrombin (1UVS) is 1.89 Å, indicating an essentially correct conformation but incorrect binding mode (Figure 3). The centroids of the crystal structure pyridine, piperidyl, and cyclohexane rings respectively, were superimposed on the centroids of the cyclohexane, piperidyl and pyridine rings respectively of the continuum solvent conformation.

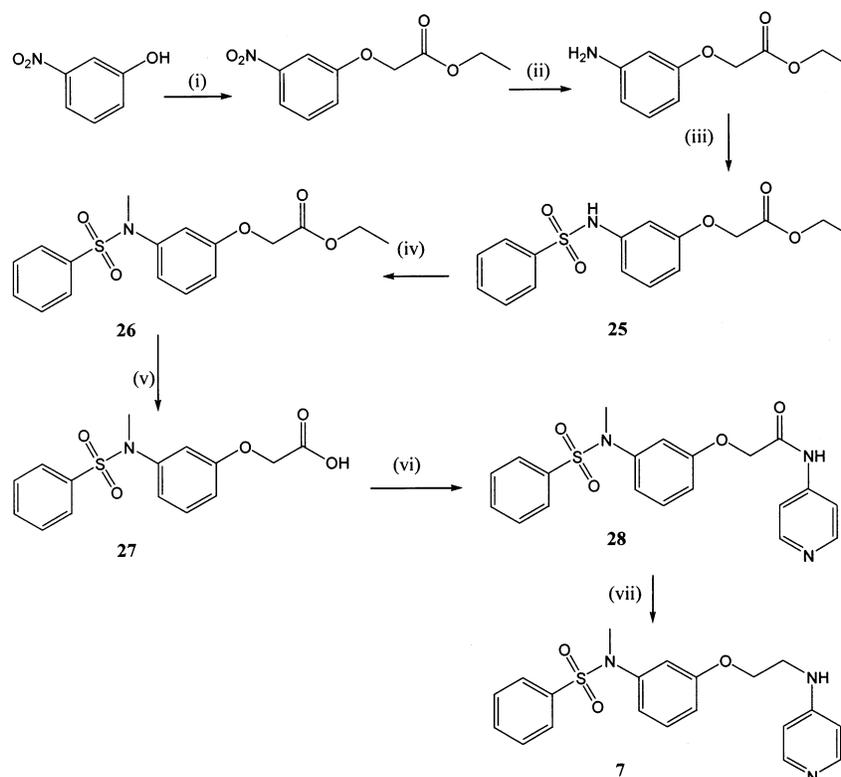
**Synthesized Test Set. Chemistry.** To facilitate parallel synthesis of the test set compounds (**7** and **15**) as well as new analogues (**20–24**, Table 4) a novel route was designed in four steps, via *N*-alkylation of the Boc protected aminopyridine with a bis-tosylate, *O*-alkylation with the appropriate aminophenol and sulfonylation (Scheme 1). The Boc-protected 4-aminopyridine **18** was readily prepared in 65% yield by treating 4-aminopyridine with di-*tert*-butyl dicarbonate in THF. *N*-Alkylation of **18**, using ethylene glycol di-*p*-tosylate, followed by nucleophilic displacement of the tosylate intermediate by 3-aminophenol, gave the intermediate **21** in 35% overall yield. The reaction of **19** with the

**Table 4.** Structures and Predicted and Experimental (exptl)  $K_i$  Values for Compounds Synthesized In-House

N <sup>o</sup>	Structure	Predicted $K_i$ ( $\mu\text{M}$ )	$K_{i\text{exptl}}$ ( $\mu\text{M}$ )
<b>20</b>		1.4	30.7 <sup>a</sup>
<b>21</b>		1.5	90 <sup>a</sup>
<b>22</b>		1.2	0.97
<b>23</b>		19	1.23
<b>24</b>		0.16	0.27
<b>15</b>		0.26	1.87
<b>7</b>		0.8	0.33

<sup>a</sup> Solubility problems.

appropriate sulfonyl chloride, in the presence of base, led to a mixture of the mono- and the di-sulfonamides **20–22** which were separated by flash chromatography. Disulfonylation was unexpected since the more electron rich 3-amino-5-methylphenol has been reported to react cleanly in good yield. Removal of the Boc group, using 50% TFA in DCM, led to the extensive decomposition of the sulfonamides, especially in the case of the benzenesulfonamides. The toluenesulfonamides **15** and **24** were obtained in very low yields (5%).

Scheme 2<sup>a</sup>

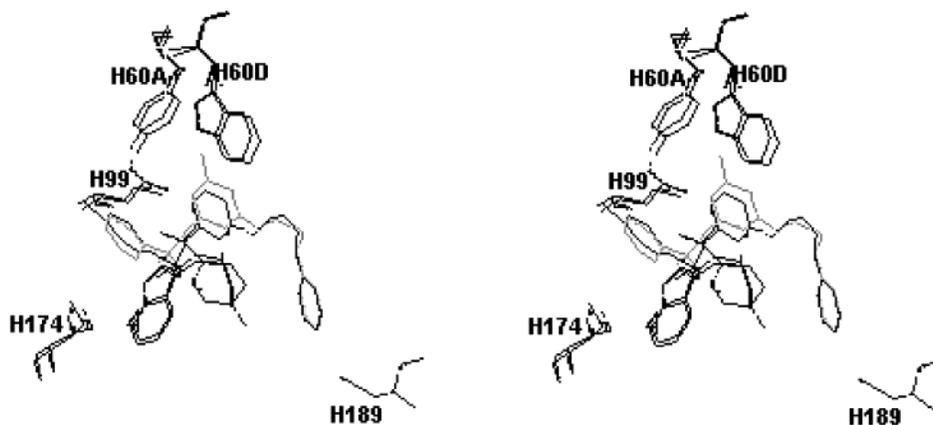
<sup>a</sup> Reagents and conditions: (i) ethyl chloroacetate, potassium carbonate, DMF; (ii) H<sub>2</sub>, Pt(IV)O<sub>2</sub>, EtOH; (iii) benzenesulfonyl chloride, pyridine, DCM; (iv) methyl iodide, potassium carbonate, DMF; (v) potassium hydroxide, MeOH; (vi) (a) thionyl chloride, (b) 4-aminopyridine, NMM, DMF; (vii) LiAlH<sub>4</sub>, THF.

Due to these poor yields, and to allow unsymmetrical substitution of the aminophenol nitrogen, a new seven-step route (Scheme 2) was investigated in which the *N*-alkyl sulfonamides were prepared via: (i) *O*-alkylation of 3-nitrophenol, (ii) reduction of the nitro group to afford the amino derivative, and (iii) reaction of the latter with benzenesulfonyl chloride, to form the desired sulfonamide **25** in 47% overall yield. To avoid disulfonylation the reaction with benzenesulfonyl chloride (one equivalent) and the mild base, pyridine, was followed by TLC and quenched at 0 °C to give **25**. *N*-methylation of the sulfonamide was carried out using methyl iodide in the presence of base in DMF to produce **26** in good yield (88%). However, hydrolysis of the ethyl ester group of the sulfonamides gave the acids in poor yields (31–54%). The acid **26** was converted into an acid chloride before being reacted with 4-aminopyridine to form the amides **30** in 53% yield. Reduction of **30** using lithium aluminum hydride in THF gave the desired sulfonamide **7** in 51% yield.

**Modeling.** Due to the difficulty of the synthesis of substituted aminophenol derivatives, this limited the diversity of the synthesized test set. Table 4 shows the structures and observed and predicted  $K_i$  values for the five new compounds synthesized in house (**20–24**) plus two additional compounds (**7** and **15**) which were prepared to provide a reference for our measured  $K_i$  values versus those in the literature. The  $K_i$  values for compounds **20** and **21** are somewhat under predicted but the poor solubility of the compounds is likely to be a contributory factor. In the pharmacophore model, the Boc group shared by compounds **20** and **21** and one tosyl group belonging to the di-tosyl compound **21** point into

the solvent. Neither the hydrogen bond donor or the S2-features are mapped by the molecules, but the S1-pocket is occupied by the pyridine ring common to them. The S4-feature and hydrophobe defined by the central aromatic ring of compound **1** are also mapped. Compounds **22** and **24** are predicted to map to the pharmacophore in a manner similar to each other, though compound **24** satisfies four out of five features as compared to only three by compound **22**, the presence of the Boc group preventing the participation of its 4-amino group in hydrogen bonding. As expected, the S2-feature is not mapped by either of these molecules. The predicted  $K_i$  values for compounds **22** and **24** are in good agreement with the observed ones, and it will be discussed below how the calculated binding modes fair with respect to the experimental data. Compound **23** is predicted to have too high an affinity and calculated to map to the pharmacophore in an inverse binding mode.

**A Comparison of the Crystal Structure Binding Modes of Compounds 1 and 24.** As in the crystal structure complex of compound **1** with thrombin, the hydrogen bonds are conserved between the 4-amino group with the Ser-214 carbonyl oxygen atom (Figure 4; N–O distance 2.77 Å for 1UVT and 3.28 Å for compound **24**, respectively), and the pyridine proton with Asp-189 (distances from the pyridyl nitrogen atom to Asp-189 OD1 and OD2 are 3.01 and 2.95 Å, respectively for 1UVT as compared to 3.1 and 2.83 Å for compound **24**). The C3 methyl group of the central aryl ring of 1UVT occupies the S2-site more fully than the unsubstituted phenyl ring of compound **24** and this results in some movement of the side-chain of Trp-60D.

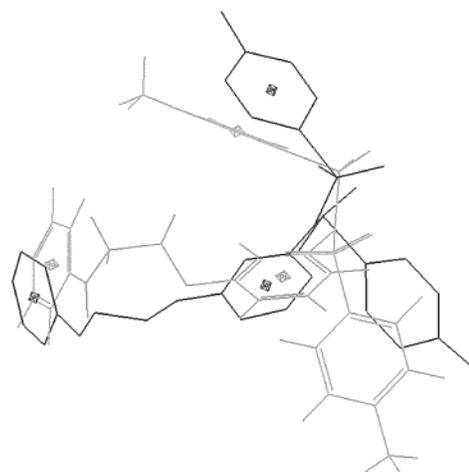


**Figure 4.** Superimposition of the crystal structures of compound **1** (bold) and compound **24** in complex with thrombin (light shade).

However, after structural alignment the RMSD between the alpha-carbons of the two crystal structure complexes is only 0.51 Å and the overall difference between the two complexes is small. One tosyl group of compound **24** engages in an edge-to-face pi-stacking interaction with Trp-215, but is displaced relative to the sulfonamide-benzene moiety of compound **1**. This is necessary to avoid a clash of the methyl group with the alpha-carbon atom of Asn-98. The separation of the Asn-98 alpha-carbon atom from the tosyl methyl of compound **24** is 4.05 Å while for the para position of the phenyl ring of compound **1** the distance is 3.48 Å. The overall conformation of the molecule is however maintained such that the dihedral angle subtended by the atoms phenylC–N–S–arylC is 64.13° in 1UVT and 83.57° for compound **24**. The second tosyl group of compound **24** is solvent exposed, and in the crystal structure is observed to form van der Waals interactions with the side chain methylenes of Glu-217 stabilizing an interaction of the Glu-217 carboxylate oxygen atom with the sulfur atom (3.07 Å for compound **24** as compared to 5.85 Å for compound **1**) of the other, inward facing, tosyl group. This interaction is possibly an artifact of the hydrophobic environment of the crystal since it is not observed in the 1UVT complex or that of the monotosyl, *N*-methyl compound **7** (unpublished results), and would require disruption of a favorable surface salt bridge between Glu-217 and Lys224 (2.90 Å). Thus, the calculated binding mode of compound **24** in complex with thrombin is consistent with the crystallographic one. Upon the basis of the superimposition of the centroids of the pyridine and central rings, and a tosyl group of compound **24** (pharmacophore and crystal structure conformation), the RMSD is 2.23 Å over 36 atoms (Figure 5). Even though no pharmacophore feature existed to describe the location of the second tosyl group, there is an obvious similarity in the spatial disposition of this moiety in the crystal structure and in the pharmacophore conformation.

### Conclusions

The suitability of the combination of a structure-based pharmacophore model and solvent generated thrombin inhibitor conformations as a 3D-QSAR method requiring limited structural information was investigated. It has long been considered that the preferred conformation of a ligand in solution is the relevant one for binding<sup>44</sup>



**Figure 5.** Superimposition of the crystal structure conformation of compound **24** (bold) extracted from the complex with thrombin and the conformation calculated to have the best fit score to the pharmacophore model (light shade).

as a close correspondence between the solution and bound conformations is entropically advantageous for the ligand.<sup>45</sup> Clustering conformers generated by an exhaustive search in continuum solvent into conformational families allowed pharmacophore mapping calculations to be more quickly performed. The quality of the 3D-QSAR model after regression with these conformational models is considered to be reduced relative to the exhaustive in continuum solvent conformational model based on the RMSD and correlation coefficient between calculated and observed affinities. However, using these two criteria the clustered in continuum solvent conformational model out-performs the 3D-QSAR model regressed with the exhaustive in vacuo conformational model, but this is largely attributable to the difference in the value of  $K_i$  calculated for one compound, compound **14**. Importantly, it was possible to rationalize differences in the binding affinities of the inhibitors based on their calculated mappings to the pharmacophore model; randomizing the experimental data led to a worse 3D-QSAR model. The binding modes of inhibitors external to the training set were well produced upon comparison with conformations when cocrystallized with thrombin. MacroModel<sup>17</sup> in continuum solvent conformational models offer an alternative to default Catalyst<sup>16</sup> conformer generation, and may result

in calculated binding modes which are more consistent with experimental data and thus provide reliable starting points for lead optimization during a medicinal chemistry program.

## Experimental Section

**Standard Method of Determination of IC<sub>50</sub> for Test Compounds.** Measurements were obtained on a Molecular Devices Corp. Thermomax plate reader using Softmax software. The chromogenic substrate S-2238 (H-D-Phe-Pip-Arg-pNA, for Thr) was obtained from Quadrant. The buffer was 0.1 M sodium phosphate; 0.2 M sodium chloride; 0.5% PEG6000 and 0.02% sodium azide. Human  $\alpha$ -thrombin was obtained from Haematologics inc.

A solution of human  $\alpha$ -thrombin (50 mL of a dilution of 1/10000 of stock at 1 mg/mL) was added after 2 min to a series of 10-fold dilutions of each inhibitor (100  $\mu$ L of dilution in buffer of 1/20 from stock in DMSO) in the presence of a fixed concentration of the appropriate chromogenic substrate (50  $\mu$ L of a dilution of 1/100 of 2.4 mg/mL stock) at 37 °C, to determine the inhibitor concentration needed to give ~50% inhibition. A series of concentrations of inhibitor either side of this approximate 50% value (of at least five different inhibitor concentrations, each point measured in duplicate) were plotted graphically to determine the exact inhibitor concentration needed to give 50% inhibition (IC<sub>50</sub>).

**Chemistry. General Methods.** <sup>1</sup>H and <sup>13</sup>C NMR were obtained on a Bruker Avance DPX400 instrument at 400.13 and 100.16 MHz, respectively, relative to TMS internal standard. Chemical shifts are given in ppm and peak multiplicities are designated as follows: s, singlet; d, doublet; t, triplet; m, multiplet. Coupling constants are given in Hz. Mass spectra were determined on a Finnigan SSQ 710C spectrometer. Elemental analyses were performed by the Elemental Analysis Department, Department of Biological and Applied Sciences of the University of North London. Results were within 0.4% of theoretical values, unless otherwise stated. All compounds exhibited NMR and MS analyses consistent with the proposed structures. All solvents and reagents were purchased from Aldrich except dimethylformamide (from Rathburn), dichloromethane, methanol and ethanol (Romil), potassium hydroxide and potassium carbonate (BDH) and hydrogen cylinders (Boc gases). The abbreviations used are as follows: DCM: dichloromethane, DMF: *N,N*-dimethylformamide, HCl: hydrochloric acid, MeCN: acetonitrile, NaOH: sodium hydroxide, NMM: *N*-methylmorpholine, TFA: trifluoroacetic acid, THF: tetrahydrofuran, TLC: thin-layer chromatography.

**4-[*N*-(*tert*-butyloxycarbonyl)amino]pyridine (18).** 4-Aminopyridine (10 g, 0.106 mol) was added slowly, in portions, to a solution of di-*tert*-butyl dicarbonate (25.5 g, 1 equiv) in THF, under argon. The reaction mixture was stirred at room temperature for 90 min, the solvent evaporated under reduced pressure and the product crystallized with diethyl ether to give a white solid (13 g, 65%, mp = 144–145 °C); <sup>1</sup>H NMR  $\delta$  1.53 (s, 9H, *t*-Bu), 7.30 (d, *J* = 6 Hz, 2H), 8.43 (d, *J* = 6 Hz, 2H); <sup>13</sup>C NMR  $\delta$  28.2, 81.8, 112.3, 150.3, 151.8, 157.3.

**3-Amino[2'-(*tert*-butoxycarbonyl)pyridin-4-yl-amino]ethoxyphenol (19).** Sodium hydride (0.11 g) was added to a stirred solution of **18** (0.5 g, 2.5 mmol) and ethylene glycol di-*p*-tosylate (0.93 g, 2.5 mmol) in dry DMF under argon. The reaction mixture was stirred for a further 7 h. 3-Aminophenol (0.27 g, 2.5 mmol) was added, followed by sodium hydride (0.08 g) and the mixture was stirred for a further 72 h. After cooling to room temperature, water (10–15 mL) was added and the mixture was evaporated under reduced pressure. The residue was partitioned between ethyl acetate and HCl (2 M). The acidic layer was exhaustively extracted with ethyl acetate, basified with NaOH pellets and extracted again with ethyl acetate. The organic phase was washed with water and brine, and dried over magnesium sulfate. After filtration of the desiccant and evaporation of the solvent a yellow oil was obtained (*R*<sub>f</sub> 0.2, ethyl acetate 100%). After purification (silica,

**Table 5.** Accurate Mass Measurements for Compounds **21**–**24**

compound	molecular formula	theoretical mass	measured mass	error (ppm)
<b>21</b>	C <sub>25</sub> H <sub>29</sub> N <sub>3</sub> O <sub>5</sub> S	484.19056	484.19056	0.1
<b>22</b>	C <sub>30</sub> H <sub>31</sub> N <sub>3</sub> O <sub>7</sub> S <sub>2</sub>	610.16815	610.16676	2.3
<b>23</b>	C <sub>24</sub> H <sub>27</sub> N <sub>3</sub> O <sub>5</sub> S	470.17495	470.17437	1.23
<b>24</b>	C <sub>27</sub> H <sub>27</sub> N <sub>3</sub> O <sub>5</sub> S <sub>2</sub>	538.14703	538.14614	1.65

ethyl acetate/hexane), 0.3 g (35%) of the pure product was collected; <sup>1</sup>H NMR  $\delta$  1.26 (m, 2H, CH<sub>2</sub>), 1.52 (s, 9H, *t*-Bu), 3.70 (br s, 2H), 4.12 (q, 2H, CH<sub>2</sub>), 6.23 (m, 2H), 6.8 (s, 1H), 6.98 (m, 1H), 7.34 (d, *J* = 5.3 Hz, 2H), 8.44 (d, *J* = 5.3 Hz, 2H); <sup>13</sup>C NMR  $\delta$  28.2, 60.4, 81.8, 102.3, 105.8, 107.5, 112.4, 130.3, 145.9, 149.9, 150.1, 151.8, 157.3; ES-MS (*m/e*) 330.0 (M+H).

**3,3-Ditosylamino[2'-(*tert*-butoxycarbonyl)pyridin-4-ylamino]ethyl]phenyl ether (20) and 3-tosylamino[2'-(*tert*-butoxycarbonyl)pyridin-4-ylamino]ethyl]phenyl ether (21).** *p*-Toluenesulfonyl chloride (0.25 g, 1.5 equiv) in DCM was added dropwise to a stirred solution of **19** (0.25 g, 0.9 mmol) and triethylamine (0.14 mL) in dry dichloromethane (5 mL). The mixture was stirred at room temperature and the reaction was followed by TLC. DCM was added and the organic phase was washed with aqueous HCl (0.1 M), water and brine, and dried over magnesium sulfate. After filtration of the desiccant and evaporation of the solvent, a yellow oil (0.39 g) was obtained. The mixture was separated by flash chromatography (silica, ethyl acetate/hexane); **20**: 0.262 g (55%); <sup>1</sup>H NMR  $\delta$  1.49 (s, 9H, *t*-Bu), 2.40 (s, 6H, 2 CH<sub>3</sub>), 4.11 (d, 4H, 2 CH<sub>2</sub>), 6.49 (m, 1H), 6.50 (m, 1H), 7.15 (m, 1H), 7.25 (m, 4H), 7.72 (m, 6H), 8.54 (d, *J* = 6.8 Hz, 2H); <sup>13</sup>C NMR  $\delta$  22.1, 28.5, 49.2, 66.4, 84.9, 103.2, 117.3, 117.4, 119.1, 125.1, 129.0, 130.0, 130.3, 135.9, 136.8, 145.6, 152.5, 158.5, 207.4; ES-MS (*m/z*) 638.3 (M+H).

**21**: 0.213 g (29%); <sup>1</sup>H NMR  $\delta$  1.26 (m, 2H, CH<sub>2</sub>), 1.49 (s, 3H, CH<sub>3</sub>), 1.53 (s, 9H, *t*-Bu), 4.1 (m, 2H, CH<sub>2</sub>), 6.54 (dd, *J* = 1.6 and 8.7 Hz, 1H), 6.66 (d, *J* = 1.6 Hz, 1H), 7.08 (d, *J* = 8.7 Hz, 1H), 7.20 (d, *J* = 8.2 Hz, 2H), 7.31 (d, *J* = 8.5 Hz, 1H), 7.33 (d, *J* = 6.4 Hz, 2H), 7.65 (d, *J* = 8.3 Hz, 2H), 8.44 (d, *J* = 6.2 Hz, 2H); <sup>13</sup>C NMR  $\delta$  21.5, 28.2, 48.7, 65.9, 81.6, 107.3, 110.7, 112.4, 113.5, 119.7, 127.2, 129.6, 130.1, 136.4, 138.3, 143.7, 145.8, 150.1, 152.0, 153.3, 158.9; ES-MS (*m/z*) 484.2 (M+H).

**3,3-Dibenzylsulfonylamino[2'-(*tert*-butoxycarbonyl)pyridin-4-ylamino]ethyl] phenyl ether (22) and 3-benzylsulfonylamino[2'-(*tert*-butoxycarbonyl)pyridin-4-ylamino] ethyl] phenyl ether (23).** These were synthesized using the same method as for **20** and **21**. **22**: 0.0037 g (14%); <sup>1</sup>H NMR  $\delta$  0.80 (m, 2H, CH<sub>2</sub>), 1.20 (s, 9H, *t*-Bu), 4.15 (m, 2H, CH<sub>2</sub>), 6.40 (s, 1H), 6.48 (d, *J* = 7.3 Hz, 1H), 6.78 (dd, *J* = 2 and 7.3 Hz, 1H), 7.20 (m, 1H), 7.50 (m, 4H), 7.62 (m, 2H), 7.81 (d, *J* = 6.1 Hz, 2H), 7.91 (d, *J* = 7 Hz, 4H), 8.57 (d, *J* = 6.1 Hz, 2H); ES-MS (*m/e*) 610.3.

**23**: 0.0046 g (17%); <sup>1</sup>H NMR  $\delta$  0.81 (m, 2H, CH<sub>2</sub>), 1.23 (s, 9H, *t*-Bu), 4.15 (m, 2H, CH<sub>2</sub>), 6.2 (dd, *J* = 1.8 and 7.1 Hz, 1H), 6.33 (d, *J* = 1.8 Hz, 1H), 6.48 (dd, *J* = 1.8 and 7.1 Hz, 1H), 6.95 (m, 1H), 7.45 (m, 2H), 7.58 (m, 1H), 7.68 (d, 2H, *J* = 5.4 Hz), 7.79 (d, 2H, *J* = 7.1 Hz), 8.56 (d, 2H, *J* = 5.4 Hz); ES-MS (*m/e*) 470.3.

**3,3-Ditosylamino[2'-(pyridin-4-ylamino)ethyl]phenyl ether (24).** A mixture of 1% anisole in TFA (2 mL) was added slowly to a stirring solution of **20** (0.1 g, 0.2 mmol) in DCM (2 mL). After 45 min at room temperature, the solvent was removed under reduced pressure and the residue obtained was dissolved in DCM, washed twice with water and brine, and dried over magnesium sulfate. After filtration of the desiccant and evaporation of the solvent, an oil was obtained which was purified by RP-HPLC (MeCN, H<sub>2</sub>O, TFA). **24** (0.005 g) was recovered. <sup>1</sup>H NMR  $\delta$  1.37 (m, 2H, CH<sub>2</sub>), 2.43 (s, 6H, 2 CH<sub>3</sub>), 4.12 (m, 2H, CH<sub>2</sub>), 6.52 (m, 2H), 7.19 (m, 2H), 7.27 (m, 4H), 7.78 (m, 6H), 8.53 (d, *J* = 7.1 Hz, 2H); ES-MS (*m/z*) 538.2 (M+H) (Table 5).

**3-Tosylamino[2'-(pyridin-4-ylamino)ethyl]phenyl ether (15).** This was synthesized from **21** following the same procedure as for **24** to give **15**. 0.075 g (95%).<sup>18</sup>

**2-(3-Nitrophenoxy)-acetic acid ethyl ester.** *meta*-Nitrophenol (5 g, 8.6 mmol) and potassium carbonate (4.97 g, 1 equiv) were stirred in DMF at room temperature while ethyl chloroacetate (3.8 mL, 1 equiv) was added. The mixture was heated under reflux at 70 °C for 5 h, and 130 °C for 48 h. After cooling to room temperature, water was added and the solution was concentrated and taken up in ethyl acetate. The organic phase was washed with NaOH (0.2 M), water, and brine and dried over magnesium sulfate. After filtering the desiccant, the solvent was evaporated under vacuum to give an oil (4.5 g, 56%); <sup>1</sup>H NMR δ 1.31 (t, 3H, CH<sub>3</sub>), 4.28 (q, 2H, CH<sub>2</sub>), 4.72 (s, 2H, CH<sub>2</sub>), 7.26 (d, *J* = 2.5 Hz, 1H), 7.46 (m, 1H), 7.72 (s, 1H), 7.87 (d, *J* = 1.9 Hz, 1H); <sup>13</sup>C NMR δ 14.2, 61.7, 65.5, 109.1, 116.8, 121.8, 130.2, 142.2, 162.6, 167.7; FAB-MS (*m/z*) 226 (M+H). Anal. (C<sub>10</sub>H<sub>11</sub>NO<sub>5</sub>) C, H, N.

**2-(3-Aminophenoxy)-acetic Acid Ethyl Ester.** A balloon of hydrogen was placed over a degassed stirred mixture of 2-(3-nitrophenoxy)-acetic acid ethyl ester (2 g, 8.9 mmol) and platinum (IV) oxide (0.10 g, 5% w/w) in ethanol (60 mL). The mixture was stirred at room-temperature overnight, filtered over Celite and concentrated down to give a colorless oil (1.66 g, 96%); <sup>1</sup>H NMR δ 1.25 (t, 3H, CH<sub>3</sub>), 3.72 (broad s, 2H, NH<sub>2</sub>), 4.26 (q, 2H, CH<sub>2</sub>), 4.57 (s, 2H, CH<sub>2</sub>), 6.26–6.34 (m, 3H), 7.05 (m, 1H); <sup>13</sup>C NMR δ 14.2, 61.3, 65.4, 102.0, 104.2, 108.9, 130.2, 147.9, 159.0, 169.1; FAB-MS (*m/z*) 196 (M+H). Anal. (C<sub>10</sub>H<sub>13</sub>NO<sub>3</sub>) C, H, N.

**2-[3-(Phenylsulfonyl)aminophenoxy]-acetic Acid Ethyl Ester (25).** Benzene sulfonyl chloride (1.51 g, 1.09 mL, 8.5 mmol) was added slowly, at 0 °C, to a stirred solution of 2-(3-aminophenoxy)-acetic acid ethyl ester (1.66 g, 8.5 mmol) and pyridine (0.69 mL, 1 equiv) in DCM (30 mL). The mixture was left to stir at room-temperature overnight, diluted with DCM and the organic phase was washed with HCl (0.2 M), water and brine and dried over magnesium sulfate. After filtering the desiccant and evaporation of the solvent, an orange solid (2.52 g, 88%) was obtained; <sup>1</sup>H NMR δ 1.28 (t, 3H, CH<sub>3</sub>), 4.26 (q, 2H, CH<sub>2</sub>), 4.55 (s, 2H, CH<sub>2</sub>), 6.61 (dd, *J* = 2.3 and 8 Hz, 1H), 6.68 (dd, *J* = 1.3 and 8 Hz, 1H), 6.74 (d, *J* = 2 Hz, 1H), 7.09 (m, 1H), 7.41 (m, 2H), 7.51 (m, 1H), 7.79 (d, *J* = 8.6 Hz, 2H); <sup>13</sup>C NMR δ 14.1, 61.5, 65.3, 107.6, 111.4, 114.3, 127.2, 129.1, 130.2, 133.1, 137.9, 138.9, 158.4, 168.8; FAB-MS (*m/z*) 326 (M+H), 358 (M+Na). Anal. (C<sub>16</sub>H<sub>17</sub>NO<sub>5</sub>S) C, H, N.

**2-[3-(Phenylsulfonyl)-*N*-methylamino-phenoxy]-acetic Acid Ethyl Ester (26).** Methyl iodide (0.42 g, 0.19 mL) was added to a mixture of 25 (1 g, 3 mmol) and potassium carbonate (0.41 g, 1 equiv) in DMF (20 mL) and the mixture was stirred for 1 h. The mixture was concentrated down and the residue was dissolved in ethyl acetate, washed with water and brine, and dried with magnesium sulfate. After filtering the desiccant and evaporating the solvents, a yellow oil (1.04 g, 86%) was obtained; <sup>1</sup>H NMR δ 1.28 (t, 3H, CH<sub>3</sub>), 3.15 (s, 3H, CH<sub>3</sub>), 4.27 (q, 2H, CH<sub>2</sub>), 4.58 (s, 2H, CH<sub>2</sub>), 6.68 (dd, *J* = 1.9 and 8 Hz, 1H), 6.73 (d, *J* = 0.6 Hz, 1H), 6.82 (dd, *J* = 2.5 and 8.3 Hz, 1H), 7.19 (m, 1H), 7.45 (m, 2H), 7.56 (m, 3H); <sup>13</sup>C NMR δ 14.2, 38.1, 61.5, 65.4, 113.3, 113.7, 119.3, 127.8, 128.6, 129.5, 132.8, 136.3, 142.7, 158.0, 168.6; FAB-MS (*m/z*) 350 (M+H).

**2-[3-(Phenylsulfonyl)-*N*-methylamino-phenoxy]-acetic Acid (27).** Powdered potassium hydroxide (0.29 g, 0.005 mol) was added to 26 (0.89 g, 2.6 mmol) and then dissolved in methanol (20 mL). The resulting mixture was stirred for 24 h. The solvent was evaporated under reduced pressure and the residue was dissolved in water and acidified to pH 4–5 with citric acid (2 M). The acidic phase was extracted with ethyl acetate, which was washed with water and brine over magnesium sulfate. The desiccant was filtered and the solvent was evaporated to give a colorless oil (0.44 g, 54%); <sup>1</sup>H NMR δ 3.16 (s, 3H, CH<sub>3</sub>), 4.64 (s, 2H, CH<sub>2</sub>), 5.05 (broad s, 1H, OH), 6.68 (d, *J* = 7.9 Hz, 1H), 6.76 (s, 1H), 6.83 (d, *J* = 8.3 Hz, 1H), 7.20 (m, 1H), 7.45 (m, 2H), 7.56 (m, 3H); <sup>13</sup>C NMR δ 38.1, 64.9, 113.5, 113.7, 119.6, 127.8, 128.8, 129.7, 132.9, 132.9, 136.2, 142.8, 157.6; FAB-MS (*m/z*) 322 (M+H).

**2-[(3-(Phenylsulfonyl)-*N*-methylamino-phenoxy)-4-pyridyl] Acetamide (28).** Thionyl chloride (0.12 mL, 1.7 mmol)

**Table 6.** Data Obtained from Crystallographic Analysis of Compound **24** with Human α-Thrombin

space group	<i>C2</i>
unit cell	<i>A</i> = 69.747 Å, <i>b</i> = 71.417 Å, <i>C</i> = 72.528 Å, β = 100.363°
resolution range Å	20–1.8 (1.85–1.80) <sup>a</sup>
no. of unique reflections	32519
overall completeness	99.1% [98.0] <sup>a</sup> (20–1.80 Å)
R-merge (%)	5.5 (32)
R-factor (R-working) (%)	18.19
R-factor (free) (%)	22.21
rms dev. bonds (Å)	0.012
rms dev angles (deg)	1.9

<sup>a</sup> Figures in brackets refer to the highest resolution bin data.

was added quickly to **27** (0.44 g, 1.4 mmol) dissolved in DMF (1 mL), and the solution heated under reflux at 70 °C. After 5 min, a solution of 4-aminopyridine (0.16 g, 1 equiv) with NMM (0.18 mL) in DMF (0.5 mL) was added slowly to the reaction mixture, which was stirred for an additional 12 h at 70 °C. After cooling to room temperature, water was added and the aqueous phase was extracted with ethyl acetate. The organic phase was washed with NaOH (0.1 M), water and brine and dried over magnesium sulfate. After filtration of the desiccant and evaporation of the solvent, an orange oil (0.29 g, 53%) was obtained; <sup>1</sup>H NMR δ 3.18 (s, 3H, CH<sub>3</sub>), 4.59 (s, 2H, CH<sub>2</sub>), 6.66 (dd, *J* = 1.7 and 7.8 Hz, 1H), 6.88 (dd, *J* = 2.4 and 8.2 Hz, 1H), 6.94 (s, 1H), 7.25 (m, 1H), 7.46 (m, 2H), 7.55–7.61 (m, 5H), 8.45 (broad s, 1H), 8.56 (dd, *J* = 1.5 and 4.2 Hz, 2H); <sup>13</sup>C NMR δ 38.1, 67.6, 113.3, 113.9, 114.6, 119.7, 127.9, 128.9, 130.0, 133.0, 136.3, 143.1, 143.8, 150.1, 156.9, 166.6; ESI-MS (*m/e*) 399.6 (M+2H).

***N*-Methyl-3-[2-(pyridin-4-ylamino)ethoxy-phenyl]benzenesulfonamide (7).** A solution of **28** (0.1 g, 0.25 mmol) in dry THF (10 mL) was added slowly to a refluxing solution of lithium aluminum hydride (0.03 g, 0.8 mmol) in THF. The reaction mixture was stirred overnight at 80 °C. After cooling to room temperature, a few drops of water were added and the solvent evaporated under reduced pressure. The residue was taken up in ethyl acetate and washed with NaOH (0.1 M), water, brine and dried over magnesium sulfate. After filtering the desiccant and evaporating the solvent, an oil (0.05 g, 51%) was obtained; <sup>1</sup>H NMR δ 2.82 (m, 2H, CH<sub>2</sub>), 3.14 (s, 3H, CH<sub>3</sub>), 3.54 (m, 2H, CH<sub>2</sub>), 4.65 (br s, NH, 1H), 6.26 (dd, *J* = 1.5 and 8.0 Hz, 1H), 6.50 (d, *J* = 4.0 Hz, 2H), 6.82 (d, *J* = 6.1 Hz, 1H), 7.09 (m, 2H), 7.19 (m, 1H), 7.45 (m, 2H), 7.56 (d, *J* = 6.1 Hz, 2H), 8.21 (d, *J* = 4.0 Hz, 2H); <sup>13</sup>C NMR δ 38.1, 42.0, 65.8, 102.6, 106.4, 107.8, 113.6, 118.5, 127.8, 128.8, 129.6, 130.1, 132.9, 149.9, 150.9, 153.3; ESI-MS (*m/e*) 385.5 (M+2H).

**Analysis of the Interaction of Compound 24 with Thrombin by Crystallography.** Human α-thrombin was obtained from Haematologic Technologies Inc., and N-Ac-Hirugen<sup>55–64</sup> was obtained from Bachem.

Crystals of the thrombin-hirugen-compound **24** complex were grown by vapor diffusion at 4 °C using the hanging drop method with 25% (w/v) PEG 8000, 0.05 M ammonium phosphate (pH 7.20) and 0.05 M sodium azide. Crystal, typically of dimensions 0.25 × 0.15 × 0.1 mm<sup>3</sup>, appeared in 4–6 weeks. Crystals were flash cooled by soaking in the crystallization buffer made up with 25% (v/v) PEG 400.

Data sets were collected using an R-axis II image plate mounted on a Rigaku RU200 anode generator to a maximum Bragg spacing of 1.8 Å, and were processed using DENZO<sup>46</sup> and Scalepack.<sup>46</sup>

Initial molecular replacement, based upon the PDB set 1H8D, was determined with AMORE.<sup>47</sup> Initial coordinates were obtained with CCP4I and the complex was refined using Refmac<sup>46</sup> with water molecules added using ARP/WARP.<sup>48</sup> Refinement converged to an R-factor of 18% (*R*<sub>free</sub> = 22%, using 5% reflections) (Table 6).

**Acknowledgment.** We would like to thank Andrew Burd for technical support and the referees for their insightful comments and suggestions.

## References

- (1) Steinmetzer, T.; Hauptmann, J.; Sturzebecher, J. *Advances In the Development of Thrombin Inhibitors. Expert Opin. Invest. Drugs* **2001**, *10*, 845–864.
- (2) Fenton, J. W., II; Ofosu, F. A.; Moon, D. G.; Maragonore, J. M. Thrombin Structure and Function: Why Thrombin Is the Primary Target for Antithrombotics. *Blood Coagul. Fibrin* **1990**, *2*, 69–75.
- (3) Huel, N. H.; Nar, H.; Priepke, H.; Ries, U.; Stassen, J.-M.; Wiene, W.; Structure-Based Design of Novel Potent Nonpeptide Thrombin Inhibitors. *J. Med. Chem.* **2002**, *45*, 1757–1766.
- (4) Gustafsson, D.; Nystrom, J.; Carlsson, S.; Bredberg, U.; Eriksson, U.; Gyzander, E.; Elg, M.; Antonsson, T.; Hoffmann, K.; Ungell, A.; Sorensen, H.; Nagard, S.; Abrahamsson, A. The Direct Thrombin Inhibitor Melagatran and Its Oral Prodrug H 375/95: Intestinal Absorption Properties, Biochemical and Pharmacodynamic Effects. *Thromb. Res.* **2001**, *101*, 171–181.
- (5) Li, C. Q.; Vindigni, A.; Sadler, J. E.; Wardell, M. R. Platelet Glycoprotein Ib alpha Binds to Thrombin Anion-Binding Exosite II Inducing Allosteric Changes in the Activity of Thrombin. *J. Biol. Chem.* **2001**, *276*, 6161–6168.
- (6) De Candia, E.; Hall, S. W.; Rutella, S.; Landolfi, R.; Andrews, R. K.; Cristofaro, R. Binding of Thrombin to Glycoprotein Ib Accelerates the Hydrolysis of PAR-1 on Intact Platelets. *J. Biol. Chem.* **2001**, *276*, 4692–4698.
- (7) Davie, E. W.; Fujikawa, K.; Kisiel, W. The Coagulation cascade: Initiation, Maintenance and Regulation. *Biochemistry* **1991**, *30*, 10363–10369.
- (8) Guinto, E. R.; Vindigni, A.; Ayala, Y. M.; Dang, Q. D.; Di Cera, E. Identification of Residues Linked to the Slow to Fast Transition of Thrombin. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 11185–11189.
- (9) Goodwin, C. A.; Deadman, J. J.; Le Bonniec, B. F.; Elgandy, S.; Kakkar, V. V.; Scully, M. F. Heparin Enhances the Catalytic Activity of des-ETW-Thrombin. *Biochem. J.* **1996**, *315*, 77–83.
- (10) Sanschagrin, P. C.; Kuhn, L. A. Cluster analysis of consensus water sites in thrombin and trypsin shows conservation between serine proteases and contribution to ligand specificity. *Protein Sci.* **1998**, *7*, 2054–2064.
- (11) Matter, H.; Defossa, E.; Heinelt, U.; Blohm, P.-M.; Schneider, D.; Müller, A.; Herok, S.; Schreuder, H.; Liesum, A.; Brachvogel, V.; Lönze, P.; Walser, A.; Al-Obeidi, F.; Wildgoose, P. Design and Quantitative Structure–Activity Relationship of 3-Amidinobenzyl-1H-indole-2-carboxamides as Potent Nonchiral, and Selective Inhibitors of Blood Coagulation Factor Xa. *J. Med. Chem.* **2002**, *45*, 2749–2769.
- (12) Jones-Hertzog, D. K.; Jorgensen, W. L. Binding Affinities for Sulfonamide Inhibitors with Human Thrombin Using Monte Carlo Simulations with a Linear Response Methodol. *J. Med. Chem.* **1997**, *40*, 1539–1549.
- (13) Bursi, R.; Grootenhuys, P. D. Comparative Molecular Field Analysis and Energy Interaction Studies of Thrombin-Inhibitor Complexes. *J. Comput. -Aided Mol. Design* **1999**, *13*, 221–232.
- (14) Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
- (15) (a) Greenidge, P. A.; Carlsson, B.; Bladh, L.-G.; Gillner, M. Pharmacophores Incorporating Numerous Excluded Volumes Defined by X-ray Crystallographic Structure in Three-Dimensional Database Searching: Application to the Thyroid Hormone Receptor. *J. Med. Chem.* **1998**, *41*, 2503–2512; (b) Gillner, M.; Greenidge, P. A. In *Pharmacophore Perception, Development and Use in Drug Design*; International University Line: La Jolla, 1999; pp 373–384.
- (16) Catalyst 4.0 Tutorials, August 1998, San Diego: Molecular Simulations Inc.
- (17) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. MacroModel-An Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics. *J. Comput. Chem.* **1990**, *11*, 440–467. MacroModel 7.2, Schrödinger, Portland, OR.
- (18) Engh, R. A.; Brandstetter, H.; Sucher, G.; Eichinger, A.; Baumann, U.; Bode, W.; Huber, R.; Poll, T.; Rudolph, R.; von der Saal, W. Enzyme Flexibility, Solvent and “Weak” Interactions Characterize Thrombin-Ligand Interactions: Implications For Drug Design. *Structure* **1996**, *4*, 1353–1362.
- (19) Schwarzl, S. J.; Tschopp, B. T.; Smith, J. C.; Fischer, S. Can the Calculation of Ligand Binding Free Energies Be Improved with Continuum Solvent Electrostatics and an Ideal-Gas Entropy Correction? *J. Comput. Chem.* **2002**, *23*, 1145–1149.
- (20) Vieth, M.; Hirst, J. D.; Brooks, C. L., III Do Active Site Conformations of Small Ligands Correspond to Low Free Energy Solution Structures? *J. Comput. -Aided Mol. Design* **1998**, *12*, 563–572.
- (21) Gund P. In: *Progress in Molecular and Subcellular Biology*, Hahn F. E., Ed., Springer-Verlag: New York, 1977; Vol. 5, pp 117–143.
- (22) Boström, J.; Norrby, P.-O.; Liljefors, T. Conformational Energy Penalties of Protein-Bound Ligands. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 383–396.
- (23) (a) Wielert-Badt S.; Lin, J. T.; Lorenz, M.; Fritz, S.; Kinne, R. K. Probing the Conformation of the Sugar Transport Inhibitor Phlorizin by 2D-NMR, Molecular Dynamics Studies, and Pharmacophore Analysis. *J. Med. Chem.* **2000**, *43*, 1692–1698. (b) Koerber S. C.; Rizo, J.; Struthers R. S.; Rivier, J. E. Consensus Bioactive Conformation of Cyclic GnRH Antagonists Defined by NMR and Molecular Modeling. *J. Med. Chem.* **2000**, *43*, 819–828.
- (24) Greenidge, P. A.; Weiser J. A Comparison of Methods for Pharmacophore Generation With the Catalyst Software and Their Use for 3D-QSAR: Application to a Set of 4-Aminopyridine Thrombin Inhibitors. *Mini-Rev. Med. Chem.* **2001**, *1*, 79–87.
- (25) Shenkin, P. S.; McDonald, D. Q. Cluster Analysis of Molecular Conformations. *J. Comput. Chem.* **1994**, *15*, 899–916. XCluster 1.7 is part of the MacroModel package, Schrödinger, Portland, OR.
- (26) Tucker, T. J.; Brady, S. F.; Lumma, W. C.; Lewis, S. D.; Gardell, S. J.; Naylor-Olsen, A. L.; Yan, Y.; Sisko, J. T.; Stauffer, K. J.; Lucaas, B. J.; Lynch, J. J.; Cook, J. J.; Stranieri, M. T.; Holahan, M. A.; Lyle, E. A.; Baskin, E. P.; Chen, I. W.; Dancheck, K. B.; Krueger, J. A.; Cooper, C. M.; Vacca, J. 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.* **1998**, *41*, 3210–3219.
- (27) von der Saal, W.; Kuczniarz, R.; Leinert, H.; Engh, R. Derivatives of 4-Amino-Pyridine As Selective Thrombin Inhibitors. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1283–1288.
- (28) Deadman, J. J.; Elgandy, S.; Goodwin, C. A.; Green, D.; Baban, J. A.; Patel, G.; Skordalakes, E.; Chino, N.; Claeson, G.; Kakkar, V. V.; Scully, M. F. Characterization of a Class of Peptide Boronates With Neutral P1 Side Chains As Highly Selective Inhibitors of Thrombin. *J. Med. Chem.* **1995**, *38*, 1511–1522.
- (29) Insight II User Guide, September 1997. San Diego: MSI, 1997.
- (30) Pauling L. *The Nature of the Chemical Bond*; Cornell University Press: New York, 1960.
- (31) Halgren, T. A. MMFF VI. MMFF94s Option for Energy Minimization Studies. *J. Comput. Chem.* **1999**, *20*, 720–729.
- (32) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.
- (33) (a) Weiser, J.; Shenkin, P. S.; Still, W. C. Approximate Atomic Surfaces from Linear Combinations of Pairwise Overlaps (LCPO). *J. Comput. Chem.* **1999**, *20*, 217–230. (b) Weiser, J.; Weiser, A. A.; Shenkin, P. S.; Still, W. C. Neighbor-List Reduction: Optimization for Computation of Molecular van der Waals and Solvent-Accessible Surface Areas. *J. Comput. Chem.* **1998**, *19*, 797–808. (c) Weiser, J.; Shenkin, P. S.; Still, W. C. Fast, Approximate Algorithm for Detection of Solvent-Inaccessible Atoms. *J. Comput. Chem.* **1999**, *20*, 586–596.
- (34) Ponder, J. W.; Richards, F. M. An Efficient Newton-like Method for Molecular Mechanics Energy Minimization of Large Molecules. *J. Comput. Chem.* **1987**, *8*, 1016–1024.
- (35) Chang, G.; Guida, W. C.; Still, W. C. An Internal Coordinate Monte Carlo Method for Searching Conformational Space. *J. Am. Chem. Soc.* **1989**, *111*, 4379–4386.
- (36) Kolossvary, I.; Guida, W. C. Low Mode Search. An Efficient, Automated Computational Method for Conformational Analysis: Application to Cyclic and Acyclic Alkanes and Cyclic Peptides. *J. Am. Chem. Soc.* **1996**, *118*, 5011–5019.
- (37) (a) Smellie, A.; Kahn, S. D.; Teig, S. L. Analysis of Conformational Coverage 1. Validation and Estimation of Coverage. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 285–294. (b) Smellie, A.; Teig, S. L.; Towbin, P. Poling: Promoting Conformational Variation. *J. Comput. Chem.* **1995**, *16*, 171–187.
- (38) Murray, C. W.; Baxter, C. A.; Frenkel, A. D.; The Sensitivity of the Results of Molecular Docking to Induced Fit Effects: Application to Thrombin, Thermolysin and Neuraminidase. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 547–562.
- (39) Coombs, G. S.; Rao, M. S.; Olson, A. J.; Dawson, P. E.; Madison, E. L. Revisiting Catalysis by Chymotrypsin Family Serine Proteases Using Peptide Substrates and Inhibitors With Unnatural Main Chains. *J. Biol. Chem.* **1999**, *274*, 24074–24079.

- (40) Weber, I. R.; Neidlein, R.; von der Saal, W.; Grams, F.; Leinhert, H.; Strein, K.; Engh, R. A.; Kuczniarz, R. Diarylsulfonamides As Selective, Non-Peptidic Thrombin Inhibitors. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1613–1618.
- (41) <http://www.msi.com/support/catalyst/hypogen.html>
- (42) Golbraikh, A.; Tropsha A. Beware of  $q^2$ . *J. Mol. Graphics and Modelling* **2002**, *20*, 269–276.
- (43) (a) Hilpert, K.; Ackermann, J.; Banner, D. W.; Gast, A.; Klaus, G.; Schmid, G.; Tschopp, T. B.; van de Waterbeemd, H. The Development of Potent and Highly Selective Thrombin Inhibitors. *Eur. J. Med. Chem. Suppl.* **1995**, *30*, 131S–138S. (b) Banner, D. W.; Hadvary, P. Crystallographic Analysis at 3.0-Å Resolution of the Binding to Human Thrombin of Four Active Site-Directed Inhibitors. *J. Biol. Chem.* **1991**, *266*, 20085–20093.
- (44) Roberts, G. C. K., *Molecular and Quantum Pharmacology*, Eds. E. D. Bergmann and B. Pullman, D. Reidel Publishing Co., Dordrecht, Holland, 1974; pp 88–90.
- (45) Lim, M. S.; Johnston, E. R.; Kettner, C. A. The solution conformation of D-Phe-Pro-containing peptides – Implications on the activity of Ac-(D)Phe-Pro-OH, a potent thrombin inhibitor. *J. Med. Chem.* **1993**, *36*, 1831–1838.
- (46) Otwinowski, Z. *In Data Collection and Processing*, 56–61 (SERC Daresbury Laboratory, 1993).
- (47) Baileys, S.. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.* **1994**, *D50*, 760–763.
- (48) Lamzin, V. S.; Wilson, K. S. Automated refinement of protein models. *Acta Crystallogr.* **1993**, *D49*, 129–147.

JM021028J