INTRODUCTION

The consideration of thermodynamic binding profiles in hit and lead optimization is increasingly relevant as a valuable analytical tool in drug discovery. Analyses of thermodynamic data of ligand binding have shown that, during late-stage optimization, improved binding is frequently achieved by enhancing the entropic component to the Gibbs free energy of binding. In many cases this results either from an appropriate rigidification of the lead scaffold in the protein-bound conformation or from the attachment of lipophilic groups of growing size to optimally fill the remaining unoccupied pockets in the binding site. Usually these strategies make the lead candidates more complex, provoking concomitant problems such as unsatisfactory bioavailability or growing risk of undesired toxicity. Accordingly, it has been recommended to start lead optimization with small, enthalpically advantaged binders, as the entropic component will be added inevitably during late-stage optimization. Therefore, methods giving reliable access to thermodynamic signatures, particularly of weak binding initial hits such as fragments, are essential to support the decision-making process, to determine which compounds to take further into lead optimization.

In this study we investigated the important drug target thrombin, a serine protease that converts soluble fibrinogen into insoluble strands of fibrin, as well as activating and catalyzing many other coagulation-related reactions. Its active site consists of three well-shaped pockets: S1, S2, and S3/4. We investigated a series of fragments binding to the S1 pocket (Figure 1). For the characterization of their binding profiles we applied isothermal titration calorimetry and used a set of more potent displacement ligands to determine their thermodynamic signature. The reference ligands used for displacement titrations address all three pockets.

In our previous study we showed that fragments binding to the S1 pocket of thrombin can be titrated either directly under so-called low c-value conditions (no sigmoidal curve) or indirectly by use of a strongly binding ligand displacing the low-affinity fragment that had already been bound to the protein in a pre-incubation step. We could show that the Gibbs free energy of binding is determined quite reliably and the obtained values are independent of the applied titration protocol, rendering the method valid for dissociation constant determinations. In contrast, the characterization of enthalpy and entropy is more difficult. Even though the displacement method achieves higher accuracy, the obtained enthalpy–entropy profile depends on the properties of the displacement ligand used in the second titration step. Although the absolute values deviate, the relative enthalpy–entropy differences determined across the experiments using the different displacement ligands revealed an unchanged signature and thus can serve as a kind of thermodynamic fingerprint for the fragments’ binding signatures.

ABSTRACT: In lead optimization, small, enthalpically advantaged fragments have been suggested to be superior, as an entropic component will be added inevitably during late-stage optimization. Determination of thermodynamic signatures of weak-binding fragments is essential to support the decision-making process, to decide which fragment to take to further optimization. High-resolution crystal structures of six fragments binding to the S1 pocket of thrombin were determined and analyzed with respect to their thermodynamic profile. The two most potent fragments exhibiting an amidine-type scaffold are not the most enthalpic binders; instead a chloro-thiophene fragment binds more enthalpically. Two chemically very similar chloro-aromatic fragments differ strongly in their potency; their binding modes are related, but the surrounding residual water network differs. The more potent one recruits a water molecule and involves Glu192 in binding, thus succeeding in firmly capping the S1 pocket. Fragments exhibiting a rather perfect solvation pattern in their binding mode also experience the highest potency.
In the current contribution we present the high-resolution crystal structures of thrombin in complex with the studied fragments and the reference ligands used for the displacement titrations. The uncovered structural details, particularly with respect to the residual solvation patterns, will be compared to the results obtained by considering the deviating thermodynamic signatures and the differences recorded by the displacement titrations.

### RESULTS

#### Crystal Structures of the Reference Ligands

Reference ligand A exhibits a predominantly enthalpic binding signature and addresses the S1 pocket of thrombin with its dichlorobenzyl moiety, the S2 pocket with a proline, and the S3/S4 pocket with a d-Phe portion (Figure 2a). At the P1 position the 2,5-dichloro-benzylamide portion forms a chlorine−π interaction with Tyr228. The 2-chlorine atom points along the C−Cl bond vector to the edge of the aromatic ring of Tyr228 at a distance of 4.9 Å. No additional water molecule can be detected in the S1 pocket. Furthermore, a hydrogen bond is formed from the NH group of this P1 group to the carbonyl oxygen of Ser214 (2.9 Å). The proline portion at P2 fills nicely the space beneath the 60-loop covering the S2 pocket. Accordingly, it prevents solvent exposure of the ligand’s central amide NH, which is hydrogen-bonded to Ser214CO. The adjacent carbonyl function (2.9 Å) and the N-terminal amino group (2.7 and 3.1 Å) of the d-Phe P3 portion form a dual β-ladder-type H-bonding motif to the backbone amide group of Gly216, which has been refined in two conformations. The Glu192 residue, spatially adjacent, also adopts two conformations with comparable occupancy pointing toward the solvent.

Reference ligand B is a less enthalpic binder and addresses the S1 pocket of thrombin with its benzamidine moiety, the S2 pocket with a proline, and the S3 pocket with a homo-Ala residue (Figure 2b). The crystal structure of reference ligand B (PDB code: 2ZGX) has been discussed previously.18 As for reference ligand A, Glu192 is present in two conformers with almost equal occupancy. Furthermore, a water molecule is found on top of the aromatic ring of Tyr228 in the S1 pocket.

Reference ligand C shows balanced enthalpic and entropic binding contributions. Similar to reference ligand B, it addresses the S1 pocket of thrombin using its benzamidine moiety (Figure 2c). It forms a salt bridge with distances both 2.9 Å to the deprotonated Asp189 at the bottom of this pocket. Furthermore, the water molecule observed in the complex with B is also present above Tyr228. In the S2 pocket, the NH of a glycine residue addresses the hydroxyl moiety of Tyr60A via a water-mediated H-bond, both with equal distances of 3.0 Å. The S3 pocket is occupied by the ligand’s d-Phe residue, and the S4 pocket is filled by a terminally attached benzylsulfonamide group. The sulfonamide moiety forms a hydrogen bond (2.9 Å) to the backbone NH of Gly219. Glu192 is disordered over several orientations, leading to a poor electron density for this side chain which has thus not been modeled into the electron density. Due to the benzylsulfonamide in the P4 position which is extending toward the S1 pocket, Glu192 is most likely pointing toward the solvent, as already observed in related structures. The binding mode of related inhibitors has also been described in a previous publication.19

#### Crystal Structures of the Fragments

Fragment 1 (Figure 2d) establishes a bifurcated hydrogen bond via its amide nitrogen to the carbonyl oxygens of Gly216 (3.3 Å) and Gly219 (3.1 Å). Both hydrogen bonds are characterized by a rather large distance, and angular relationships indicate strong deviations from ideal linearity (N-H−O=C, with 128° and 130°). The pyridine nitrogen of 1 forms an additional polar contact to Gly219 (2.9 Å). The mutual arrangement of the involved groups does not suggest the formation of an energetically favorable H-bond (N−H−O=C, as it deviates strongly from linearity (126°). In contrast to fragments 2−6, no further direct interactions with the protein in the deeply buried S1 pocket could be recognized. Fragment 1 could be refined with full occupancy at a resolution of 1.18 Å. Glu192 is observed in one single conformation and orients toward the solvent.

Fragment 2 forms a chlorine−π interaction with Tyr228 deep in the S1 pocket where the chlorine atom of the chlorothiophene points toward Tyr228 (Figure 2e). This chlorine−π interaction in protein−ligand complexes is characterized by an “edge-on” geometry, as the carbon−chlorine bond vector points toward a C−C bond of the adjacent aromatic portion.20 Furthermore, hydrogen bonds are formed between the carboxamide group of 2 and Glu192 (2.9 Å) as well as Gly219 (3.1 Å). One of the carboxylate oxygens of Glu192 is...
Figure 2. (a) Reference ligands A (light blue), (b) B (turquoise), and (c) C (cyan) and fragments (d) 1 (white), (e) 2 (green), (f) 3 (orange), (g) 4 (yellow), (h) 5 (magenta), and (i) 6 (red) bound to the active site of thrombin. (j) The apo-structure of thrombin (2UUF, light green) is shown. The protein is indicated by its solvent-accessible surface (white). Oxygen, nitrogen, and sulfur atoms are indicated in red, blue, and yellow, respectively. In (a) also the specificity pockets are assigned. Water molecules are displayed as spheres in the color of the respective ligand. Electron density difference maps ($F_o - F_c$) are shown at the 2$\sigma$ level for fragments 1 and 6 and at 3$\sigma$ for all other ligands as green meshes, respectively.
placed in a position suited to address the carboxamide nitrogen of the fragment, allowing to establish a favorable H-bonding geometry (162° to Gly219 and 173° to Glu192). As a consequence, Glu192 adopts one definite conformation. This contact, along with the support from reasonably refined B-values, suggests the unambiguous assignment of O and N to the electron density of the carboxamide group. In addition, an interstitial water molecule (red circle, Figure 2e) mediating hydrogen-bonding contacts between the carboxamide oxygen of 2 and the backbone of amino acids Cys191, Gly193 (amino acids not shown in Figure 2e), and both alternative conformations (occupancy 0.42 and 0.58) of Ser195 is observed.

The amidino moiety of fragment 3 (Figure 2f) is accommodated in the S1 pocket of thrombin and forms a salt bridge with distances of 2.9 and 2.8 Å to the most likely deprotonated Asp189 at the bottom of this pocket. Planes defined by the atoms of either the ligand’s amidino and the carboxylate group of Asp189 are twisted by 14.5° from coplanarity, which is assumed to represent the optimum for this contact. This contrasts the binding geometry observed for 4 (see below). In anti-direction the amidino nitrogens of 3 form H-bonds to the backbone carbonyl oxygen of Ala190 (3.1 Å) and Gly219 (2.9 Å). A water molecule mediating an H-bond to the carbonyl oxygen of Phe227 (not shown in Figure 2f) can be detected, and Glu192 is observed in one conformation which partly seals the S1 pocket from above. Beyond the phenyl ring

Figure 3. (a) Superposition of 3 and 4; (b) 2 and 5; (c) 5 and 6; (d) 4, C, and apo; and (e) 1, 2, 5, 6, A, and apo. Color code similar to Figure 2.
of 3 toward the solvent environment an elaborate water network is detected, consisting of seven water molecules and involving the carboxylate group of Glu192.

Via its amidino group, the amidinopiperidine fragment 4 (Figure 2g) forms a salt bridge with comparable distances as in 3 (2.9 and 2.9 Å) to the deprotonated Asp189 at the bottom of the S1 pocket. Here the twist angle between planes through the amidino and carboxylate groups approaches more closely the optimal coplanarity compared to 3 (3.6°). Further H-bonds can be observed from the amidino nitrogens to the backbone oxygen atoms of Ala190 (3.3 Å) and Gly219 (2.9 Å). The H-bond to Ala190 is longer compared to 3 (3.1 Å). The ligand’s piperidine ring adopts a chair conformation fitting well below one of the two observed Glu192 conformers which refining to almost equal occupancies of 49% and 51%. Whereas one conformer is integrated into the water network capping the S1 pocket, the second conformer is oriented toward the solvent. None of the conformers is involved in a direct contact to other residues of the protein. Remarkably, both show an orientation and accordingly torsion angles that deviate from the other structures. Similar to 3, the corresponding water-mediated H-bond to Phe227 could be detected, and the water molecule is located also at the same position as in the complex with 3. The generated water network exhibits similar features as the one observed for 3, except that Glu192 occurs in a second conformation and only the first one is involved in the network. The water network above the S1 pocket and 3 and 4 is largely conserved (Figure 3a).

Fragment 5 (resolution: 1.12 Å) refines to an occupancy of 0.80. It also forms a chlorine−π interaction to Tyr228 in the S1 pocket, and the carbon−chlorine bond vector points toward a C=C bond of the aromatic ring of Tyr228 (Figure 2h). In this structure Glu192 is disordered and no interactions to 5 can be detected. A hydrogen bond is formed between the carbonyl group of the fragment’s amide group and one conformer (occupancy 0.54) of the hydroxy methylene side chain of the catalytic Ser195 (2.7 Å). Here, the assignment of N and O to the density of the fragment’s carboxamide groups is not unequivocally defined by the H-bonding topology, but the B-factor refinement supports our allocation. In addition, water-mediated hydrogen bonds are formed via the NH group of the amide function of 5 and the carbonyl oxygens of the backbone amide groups of residues Gly216 and Gly219.

For fragment 6 (resolution: 1.24 Å) a chlorine−π interaction to Tyr228 is also observed in the S1 pocket since the chlorine atom of the para-chloro-benzyl portion points toward Tyr228 (Figure 2i). An additional hydrogen bond is formed via the NH group of the sulfamide unit to the carbonyl group of Gly219 (3.0 Å). Glu192 is oriented toward the solvent, as similarly observed for fragment 1. In addition, an interstitial water molecule mediates a hydrogen bond between one oxygen of the sulfamide group of 6 and the backbone amide groups of Gly193 and Ser195. Furthermore, the terminal dimethyl amino group of the sulfamide is involved in a water-mediated contact to the peptide backbone of Gly216. Fragment 6 refines to an occupancy of 0.75. In the difference electron density, weak density signals attributable to several water molecules at the position of the phenyl ring of 6 became visible, suggesting partial occupancy of 6 and corresponding to the water structure of the apo-protein. Due to the low occupancy, these water molecules were not included in the refinement model.

One water molecule (O1279; apo-structure 2UUF; red circle in Figure 2j) is known to be important for the binding of different P1 head groups in the S1 binding pocket of thrombin. It is found in the benzamidine (B, C, 3) and amidino piperidine (4) complexes (Figure 3d). The location of this water molecule seems to be optimal in the apo-protein, as it builds up a network with other water molecules to interact with Asp189. Furthermore, it forms a short contact to water molecules placed on top of the aromatic ring of Tyr228. Interestingly, the chlorine substituents of 2, 5, 6, and A and also the methoxy moiety of 1 are located close to the latter water position above the aromatic ring of Tyr228 and displace it from the S1 pocket (Figure 3e). With the amidine-type ligands this water molecule remains in the S1 pocket but it is pushed into a slightly more distal position compared to the apo-protein.

**Influence of Bound Hirudin Peptide on the Thermodynamic Signature.** All crystal structures presented in this study originate from crystals that were prepared in the presence of a 12-residue-long peptide derived from the 66 amino acid natural product inhibitor hirudin in order to block the exo-site of thrombin. Even though the binding site of this peptide is rather remote from the active site where all fragments and the reference ligands bind, it cannot be excluded that the binding of the peptide has some influence on their binding properties. In contrast to the crystallographic analysis, all titration experiments discussed in our previous communication were performed in the absence of the peptide. Accordingly, to study the influence of the hirudin peptide used in the crystallization procedure, we titrated the peptide directly to thrombin and performed additional titrations with the three displacement ligands A, B, and C in the presence of the hirudin peptide in order to compare them with the thermodynamic data obtained in the presence and absence of the exo-site binder.

A direct titration of the hirudin peptide to thrombin revealed no significant enthalpic signal that could be evaluated, likely due to a predominantly entropic binding profile. Subsequently, we titrated the three displacement ligands to the protein preincubated with the hirudin peptide. The measured binding enthalpies for ligands A (with hirudin ΔH = −47.0 ± 1.1 kJ/mol, without hirudin ΔH = −46.5 ± 0.3 kJ/mol), B (with hirudin ΔH = −39.9 ± 0.7 kJ/mol, without hirudin ΔH = −40.2 ± 0.6 kJ/mol), and C (with hirudin ΔH = −21.0 ± 0.2 kJ/mol, without hirudin ΔH = −22.4 ± 0.9 kJ/mol) show almost identical thermodynamic signatures within experimental accuracy (Supporting Information, Figure A). Hence, the presence of the hirudin peptide seems to have no measurable impact on the thermodynamic properties of the binding of our active-site inhibitors.

**Analysis of the Conformational Flexibility of Glu192.** A survey of a large number of high-resolution thrombin structures in the Protein Data Bank (PDB) shows that Glu192 is quite flexible and adopts multiple conformations in the various crystal structures.

We analyzed the different orientations of glutamic acid 192 (Glu192) in 174 thrombin crystal structures. All 174 thrombin structures with a resolution better than 2 Å were taken from the PDB, and extended by our eight new structures discussed in this paper. The complex with B has already been deposited and is comprised in the retrieved PDB set. A list of all considered PDB codes can be found in the Supporting Information.

In total, 13 structures were removed from the data set due to major deviations in the backbone, or cases missing Glu192 side chains: 1HAG, 2AQO, 2AFQ, 2GP9, 2Z03, 3BEI, 3DUX, 3HK3, 3JZ1, 3K65, 3SQE, SAFY (structure S), and SAFZ.
(structure C). Only 12 structures of the sample set showed a second alternative conformations of the Glu192 side chain; one had even a third orientation. In total our analysis covers 187 side chain conformers.

A distance matrix was calculated on the basis of the root-mean-square deviation (rmsd) of all side chain atoms. For visualization, a hierarchical cluster analysis using the average linkage method was performed. The distance criterion for the separation into distinct clusters was a dissimilarity of 1.7 Å. Supporting Information, Figure B shows the dendrogram used for the hierarchical clustering of the 187 Glu192 side chain entries. Figure 4 displays the distribution of the Glu192 side chain orientations in stick representation and colored according to their cluster affiliation.

The Glu192 side chain found in complex 1A4W with a thiazole-containing inhibitor probing the S1′ pocket is structurally remote to all other entries and remains as a singleton (shown in orange). It appears puzzling that a rather flexible and can adopt, depending on the bound ligand, different conformations.

**DISCUSSION**

The thermodynamic characterization of the reference ligands A–C and the four fragments 1–4 has been described in our previous contribution. 11,13 Fragments 5 and 6 are characterized by too low affinities to record their thermodynamic signature and accurate affinity data. According to a fluorescence-based assay, 13 we estimate that their binding must be about 10 mM. Thus, both are by 5–100-fold less potent than the other fragments studied. A summary of the binding data of A–C and 1–4 is given in Figure 5.

Independent of the selected titration strategy (direct/displacement) or the chosen reference ligand, the free energy of binding is determined with rather small deviations in all cases, rendering ITC as a reliable method to determine ΔG and dissociation constants. The characterization of the enthalpy/entropy partitioning is more difficult and seems to be dependent on the applied titration protocol. It has to be kept in mind that in ITC enthalpy and entropy do not result from independent measurements; they are calculated as the numerical difference between ΔG and ΔH. We therefore have to consider with respect to the relative comparisons only one of the two properties, ΔH or −TΔS. To reduce the influence of possibly superimposed protonation effects, all titration experiments were performed in pyrophosphate buffer which exhibits a small heat of ionization. Multiple titrations of ligands to thrombin have indicated that His57 generally releases approximately 0.3–0.5 mol of protons upon ligand binding. 6,19

For ligands with a scaffold like A and B, entrapment of a similar molar quantity of protons has been recorded, compensating the proton release from His57. Thus, for the studied fragments some small uncertainties of the thermodynamic profile can be given owing to superimposed and not corrected protonation effects.

Interestingly, quite consistent thermodynamic profiles are found for fragment 2 by all titration protocols. In the case of 1 this conclusion is difficult to make, as the accuracy of the measurements is limited due to the low potency of this fragment. Thus, differences in the profiles resulting from the deviating protocols have to be discussed with some care. Nonetheless, the data suggest that this ligand is a predominantly enthalpic binder. Considering the direct titration, both amidine-type fragments exhibit similar profiles, although 4 is more potent owing to both an enthalpic and entropic advantage over 3. The displacement titrations provide a more complex picture. Here the reference ligand A featuring a chloro-aromatic P1 portion reveals for both fragments 3 and 4 an entropically more favored profile, whereas the benzamidine derivatives B and C result in stronger enthalpically favored binding. The absolute values differ for both fragments; however, the relative entropic advantage of using A compared to the enthalpic components using B and C is similar (∼8–10 kJ/mol) (Figure 5b).

**Methoxy Fragment 1.** The methoxy group of 1 falls very close to the location of the chlorine substituent found in A, 2, 5, and 6 (Figure 3c). In all ligands the attached substituents are able to displace the water molecule located at this position on top of the aromatic moieties of Tyr228 in the apo-protein (PDB codes: 2UUF, 3D49) and the two amidino complexes. The methoxy group attached to fragment 1 is also found in the marketed factor Xa inhibitor Apixaban, rendering this substituent as competent to collect sufficient affinity contribution in the S1 pocket of serine proteases. 25 Fragment 1 has been refined with full occupancy and results in a B-factor ratio between fragment and protein (Bfragment/Bprotein = 1.24) which is the highest within the studied series. This indicates that the occupancy is below 100% and/or this fragment exhibits a significant residual mobility in the pocket due to less tight fixation. This observation would suggest a reduced enthalpic signature in favor of an enhanced entropic contribution. At least compared to 2 that also displaces the named water molecule (see below), it is indeed a less enthalpic binder.

**Amidine Fragments 3 and 4.** The conformationally more flexible amidinopiperidine 4 loses a larger amount of degrees of
freedom upon binding as it is restricted to only one distinct chair conformation in the S1 pocket. This should be entropically unfavorable. In comparison, the nearly planar scaffold of benzamidine allows at most some wiggling motions of the amidino group with respect to the adjacent aromatic portion, and part of this mobility will still be experienced at the binding site, which should be entropically beneficial. Additionally, the binding of 4 allows establishment of two conformations of Glu192, whereas for 3 only one conformation could be observed. Possibly the latter effect overwhelms and provides an explanation for the higher entropic binding contribution of 4 (cf. below 2 and 5). It is tempting to speculate whether this observation results in a weaker salt bridge formed to Asp189 in the case of 3, particularly as 3 is also the enthalpically less favorable binder compared to 4. Overall 4 shows more potent binding to thrombin, which stimulated the researchers at Roche to use the amidinopiperidine moiety as a P1 headgroup in the development compound napsagatran.23 Taking the data from direct titrations (as they are not further affected by superimposed implications of the displacement ligands) as the basis to compare the thermodynamic signatures of 4 over 3, the affinity advantage of 4 seems to result from both enthalpic and entropic benefits.

Interestingly, a direct comparison of the water networks formed above the S1 pocket shows no substantial differences (Figure 3a) and will hardly contribute to the differences in the thermodynamic profiles of 3 and 4.

Chlorine-Substituted Fragments 2, 5, and 6. The less potent fragment 2 shows interestingly an even lower $B_{\text{fragment}}/B_{\text{protein}}$
Table 1. Data Collection and Refinement Statistics

<table>
<thead>
<tr>
<th>compound (PDB code)</th>
<th>1 (5APF)</th>
<th>2 (4UD9)</th>
<th>3 (4UEH)</th>
<th>4 (4UE7)</th>
<th>5 (5AHG)</th>
<th>6 (5AHJ)</th>
<th>A (4UDW)</th>
<th>C (5AFZ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wavelength (Å)</td>
<td>0.976</td>
<td>0.976</td>
<td>0.9184</td>
<td>0.9184</td>
<td>0.9184</td>
<td>0.976</td>
<td>0.976</td>
<td>0.895</td>
</tr>
</tbody>
</table>
| resolution range (Å)
  (r/ε)               | 70.96−1.18| 35.68−1.12| 35.76−1.16| 19.56−1.13| 18.45−1.12| 43.41−1.24| 71.13−1.16| 21.47−1.53|
| space group         | C 1 2 1  | C 1 2 1  | C 1 2 1  | C 1 2 1  | C 1 2 1  | C 1 2 1  | C 1 2 1  | C 1 2 1  |
| unit cell a, b, c (Å)| 70.41 73.21| 70.41 73.21| 70.41 73.21| 70.71 75.24| 70.71 75.26| 70.71 73.22| 69.6 71.23| 70.71 72.4 |
| α, β, γ (deg)       | 90 100.490| 90 100.490| 90 100.590| 90 100.590| 90 100.590| 90 100.590| 90 100.590| 90 100.590|
| total reflections    | 497434   | 804825   | 444822   | 392964   | 417074   | 326876   | 506370   | 201127   |
| unique reflections   | 108698   | 127839   | 116712   | 130988   | 134540   | 95099 (8909) | 116096   | 53119 (5257) |
| multiplicity         | 4.6 (4.1)| 6.3 (4.2)| 3.8 (3.4)| 3.0 (2.7)| 3.1 (2.8)| 3.4 (3.3)| 4.4 (3.5)| 3.8 (3.7)|
| completeness (%)     | 95.1 (86.1)| 96.7 (76.3)| 95.7 (84.3)| 99.5 (99.5)| 99.3 (98.5)| 96.1 (90.7)| 97.1 (87.9)| 99.8 (99.4)|
| mean 1/o(R)         | 196.2 (2.6)| 218.3 (3.1)| 17.7 (2.5)| 11.1 (2.5)| 11.5 (2.5)| 14.2 (2.5)| 18.5 (2.6)| 17.5 (2.5)|
| Wilson B-factor      | 13.4     | 12.7     | 13.6     | 10.7     | 11.4     | 13.3     | 13.2     | 17.0     |
| R-merge              | 0.033 (0.565) | 0.036 (0.414) | 0.032 (0.400) | 0.038 (0.390) | 0.040 (0.396) | 0.043 (0.528) | 0.025 (0.472) | 0.048 (0.527) |
| R-work               | 0.121 (0.223) | 0.125 (0.186) | 0.121 (0.204) | 0.124 (0.189) | 0.122 (0.186) | 0.123 (0.229) | 0.124 (0.214) | 0.145 (0.176) |
| R-free               | 0.137 (0.247) | 0.141 (0.197) | 0.138 (0.214) | 0.139 (0.200) | 0.138 (0.190) | 0.144 (0.250) | 0.141 (0.230) | 0.174 (0.191) |
| no. of non-hydrogen 
  atoms                          | 2778      | 2815      | 2810      | 2810      | 2780      | 2659      | 2808      | 2645      |
| macromolecules       | 2434      | 2465      | 2447      | 2434      | 2416      | 2294      | 2456      | 2383      |
| ligands              | 58        | 58        | 22        | 55        | 37        | 61        | 41        | 43        |
| water                | 280       | 291       | 331       | 321       | 325       | 300       | 309       | 218       |
| protein residues     | 289       | 294       | 289       | 290       | 289       | 280       | 290       | 287       |
| RMS bonds (Å)        | 0.011     | 0.009     | 0.01      | 0.01      | 0.01      | 0.01      | 0.01      | 0.006     |
| RMS angles (deg)     | 1.43      | 1.38      | 1.46      | 1.45      | 1.43      | 1.51      | 1.51      | 1.54      |
| Ramachandran favored 
  (%)                       | 97        | 97        | 97        | 97        | 97        | 97        | 97        | 98        |
| Ramachandran outliers 
  (%)                      | 0         | 0         | 0         | 0         | 0         | 0         | 0         | 0         |
| Ramachandran allowed 
  (%)                      | 3         | 2         | 2         | 3         | 3         | 3         | 3         | 2         |
| average B-factor (Å²) | 20.6      | 20.1      | 22.0      | 17.9      | 18.9      | 19.9      | 21.0      | 25.0      |
| macromolecules       | 19.2      | 18.9      | 20.3      | 16.0      | 17.1      | 17.9      | 19.5      | 24.3      |
| ligands              | 26.8      | 27.8      | 22.9      | 20.0      | 23.1      | 28.8      | 16.8      | 26.3      |
| solvent              | 31.8      | 28.4      | 33.8      | 30.9      | 31.8      | 32.8      | 33.3      | 32.0      |
| Bfactor²/Bproteins   | 1.24      | 0.81      | 0.87      | 0.85      | 0.98      | 1.18      | –         | –         |
| coordinate errors (Å) | 0.09     | 0.07      | 0.07      | 0.08      | 0.11      | 0.08      | 0.12      | 0.12      |

aStatistics for the highest-resolution shell are shown in parentheses. bCalculated by MOLPROBITY. cMaximum-likelihood based.
and C all show one residual water molecule in S1, occupying the position of the chloride atom in 2, 5, 6, and A (Figure 3d,e). This water molecule, found on top of the aromatic ring of Tyr228, is slightly shifted in the complexes compared to its position adopted in the apo-structure. The desolvation cost required to shed the hydration shell around the charged amidino group in 3, 4, B, and C will be significantly larger than for those with the chloro-aromatic P1 group. Desolvation is predominantly enthalpically unfavorable and will therefore reduce the enthalpic binding signal. \(^{12}\) Remarkably, the chlorothiophene anchor used in 2 has also been applied as P1 headgroup in the marketed factor Xa inhibitor rivaroxaban. \(^{13}\,14\)

**Deviating Thermodynamic Profiles Obtained Using Different Reference Ligands.** We suppose that we record with these displacement titration experiments another case of non-additivity in protein–ligand binding. In case every individual step in a displacement titration procedure would be at equilibrium, the thermodynamic signature obtained for the displacement of the bound fragments should be independent of the properties of the applied stronger binding reference ligand.

A displacement titration starts by incubating the protein with the probe fragment. Subsequently, the selected reference ligand is added and repels the fragment from the pocket. If immediately the binding pocket would return to the unperturbed solvation pattern observed for the apo-protein, the displacement titration should reveal—independent of the used displacement ligand—the same results. This is due to the fact that in the evaluation the results of the titration using the reference ligand alone are subtracted from the combined titration protocol. However, within experimental accuracy this is not the case, we obtain deviations using different reference ligands. This suggests that the protein does not return intermediately to the unperturbed solvation structure seen in the apo-protein.

The chlorothiophene fragment 2 displaces all four well-ordered water molecules found in the S1 pocket of the apo-protein. Displacement by the chloro-aromatic reference ligand A occurs without changing the solvation pattern in the S1 pocket. Displacement by the two benzamidine-type reference ligands B and C involves, however, the displacement of 2 from the fully desolvated pocket and the benzamidine reference ligand enters the site together with a water molecule. Thus, the latter ligand does not fully release its solvation shell. This situation is different if the amidine-type fragments 3 and 4 are displaced from the S1 pocket. They incorporate already the water molecule on top of Tyr228. This has to be repelled once the chloro-aromatic reference A accommodates the pocket. This release is likely combined with an entropic advantage. In contrast the displacement of 3 and 4 by the benzamidine-type reference ligand B and C does not change the solvation pattern in S1, the water molecule on top of Tyr228 remains. Obviously, this shifts the thermodynamic signature relative to the displacement with A toward a more enthalpically favored process of about 8–10 kJ/mol. It remains interesting to see that the displacement of 3 and 4 obtains the same relative changes in the enthalpy/entropy inventory. But compared on an absolute scale a difference is experienced. Obviously, additional effects are overlaid to the process, resulting from further changes of the protein and/or the solvation structure, possibly involving the conformational transitions of Glu192. This underlines again that thermodynamic signatures should not be compared on an absolute scale but only relative to each other. \(^{12}\) The example shows that the mutual displacement of ligands in binding pockets can follow a rather complex and not necessarily an additive protocol. Instead cooperativity resulting from differences in the solvation pattern can be given.

**CONCLUSION**

Fragment-based lead discovery (FBLD) has emerged as a promising method to discover starting points for novel leads in a medicinal chemistry program. The obtained hits are usually characterized in terms of their binding affinity, and binding modes are determined by crystallography. However, as they are very weak binders, particularly their energetic ranking remains difficult, and computational methods often fail to provide reliable information. Subsequently, the difficult decision has to be made: which fragment to select for further optimization. Accordingly, current research seeks for criteria to support this decision-making process. Particularly, the thermodynamic properties have been hypothesized as an important gauge, and enthalpic ligand efficiency has been suggested as important. \(^{15}\)

To provide some insight into the sustainability of these criteria, we determined the high-resolution crystal structures of six fragments binding to the S1 pocket of thrombin and analyzed them with respect to their thermodynamic binding profile. We selected fragments which exhibit the well-known P1 head groups, successfully used in serine protease inhibitors such as dabigatran, napsagatran, rivaroxaban, or apixaban. Two rather weak binding fragments (5, 6) address the S1 pocket via their chloro-aromatic portion; the remaining fragment moieties do not succeed in establishing strong interactions with the enzyme. As a consequence, they show significant residual mobility, and their weak potency makes a detailed thermodynamic characterization impossible. Another chlorothiophene fragment (2) is similarly recognized in the S1 pocket through a chloride–π interaction with Tyr228. However, additional strong interactions beyond the S1 pocket and an extensive water network capping the fragment’s binding position turn this fragment into the most enthalpic binder of the series and renders 2 a much better binder compared to the highly isostructural 5. Part of the enthalpic advantage of 2 is compensated by an unfavorable positive entropic contribution, making this fragment the third most potent of the series. A fourth fragment (1) exhibiting a methoxy aromatic P1 headgroup binds with less affinity to the protein; however, its binding is also dominated by enthalpy. In addition to its S1 binding, this fragment experiences several polar contacts with the protein beyond the S1 pocket. Importantly enough, all four fragments displace four well-defined water molecules found in the S1 pocket of the apo-structure of thrombin, particularly including the water molecule located on top of the aromatic ring of Tyr228. The two most potent fragments studied are those of the amidino-type molecules 3 (benzamidine) and 4 (amidinopiperidine). The latter fragment turns out to be the best binder of the series owing to an enthalpic and entropic advantage over 3. For both charged fragments a huge enthalpic desolvation price will be required which should be much larger than for the other four fragments. \(^{12}\) This cost will, however, be overcome by the strong salt bridge formed to Asp189. Compared to the apo-structure, the water molecule on top of Tyr228 remains in the pocket, while it is slightly shifted out of the apo-position. In addition, both fragments allow for a perfect water network capping the S1 pocket even though they are not involved in any further polar contacts.

DOI: 10.1021/acs.jmedchem.5b00812

J. Med. Chem. XXX, XXX, XXX–XXX
The side chain of Glu192, which exhibits large conformational flexibility across the large variety of thrombin crystal structures, adopts different configurations in the six fragment complexes, partly involving its carboxylate group into interactions with the fragments. Supposedly, also this residue, in addition to the residual solvation structure, has an impact on the thermodynamic signature of the fragments. Remarkably, the fragments exhibiting a rather perfect solvation pattern in the adopted binding pose also experience the highest potency. Likely in the case of fragment binding, the influence of the quality and completeness of the water network wrapping around the bound fragment is even more determining for the thermodynamic signature than in the case of drug-size ligand binding.12

This observation is further emphasized by the fact that displacement titrations using different reference ligands produce deviating absolute signatures for the fragments being replaced. The displacement of amide-based fragments by a benzamidine-type ligand does not alter the S1 solvation pattern, whereas a chloro-aromatic ligand used for displacement will induce a change of the solvation pattern in the S1 pocket as the water on top of Tyr228 is repelled. These changes are superimposed to the overall titration and have to be regarded. Overall we can conclude that at least a semi-quantitative estimate on the thermodynamic signature of the fragments is possible; however, a detailed control of the involved structural changes is mandatory, particular with respect to the residual solvation structure. Only then does the assignment of a fragment to be a more enthalpic or entropic starting point for subsequent lead optimization make sense. However, for such weak binders, a modulation of the solvation structure can easily shift the thermodynamic signature from “more enthalpic” to “more entropic”. One caveat cannot be concealed. The two isostructural chloro-aromatic fragments 2 and 5 could be prospective starting points for lead optimization, and thermodynamics clearly renders 2 the more promising candidate. Experience in thrombin research, however, shows that both the P1 chloro-thiophene and the m-chloro-benzyl group are promising; thus, a weak starting point such as 5 can easily gain potency upon optimization and turn into a “well-suited” one with respect to the final optimization product. The multitude of fragments detected in a lead discovery campaign have to be seen as a kind of idea generator and information source to map active-site properties of the studied target protein.

## MATERIALS AND METHODS

### ITC Measurements. ITC experiments were performed using an ITC200 system from GE Healthcare (Northampton, MA, USA). Details about the applied measuring protocols have been described elsewhere.13

Thrombin was obtained from CSL Behring (Marburg, Germany) and purified from Beriplast. Thrombin was extracted by dialysis using an experimental buffer of 50 mM sodium tetraborate, 100 mM NaCl, 0.1% polyethylene glycol 8000 at pH 7.8. Subsequently, the protein could be used for all titration experiments. The protein concentration was measured by absorbance at 280 nm using a NanoDrop 2000c spectrophotometer from Thermo Scientific.

For studying the influence of bound hirudin, 250 μM hirudin peptide (54–65) (sulfated) from Bachem (Switzerland) in TPSS buffer was preincubated with 27.5 μM thrombin and afterward titrated. Displacement ligands from 50 mM DMSO stock solutions were diluted to 0.25 mM in ITC buffer and subsequently the DMSO concentration in all solutions was adjusted to 3%. All titrations were performed in duplicate in order to save hirudin material.

All ITC experiments were performed at 25 °C with a reference power of 5 kcal/s after a stable baseline had been reached. The pre-titration delay was set to 300 s. For titrations, ligand injections of 0.3 μL (to prevent artifacts arising from small syringe leakages or air in the syringe) were followed by 19–27 injections of 1.5–2.0 μL, with at least a 180 s interval between subsequent injections. Data integration and evaluation was performed with Ntipic29 and Sedphat.30

---

**X-ray Data. Crystallization and Soaking.** Human α-thrombin (from Enzyme Research Laboratories, South Bend, IN, USA) was dissolved in the crystallization buffer (20 mM NaH2PO4, 350 mM NaCl, 2 mM benzamidine, pH 7.5) at 10 mg/mL. The hirudin fragment Hirudin (54–65) (sulfated) obtained from Bachem (Bubendorf, Switzerland) was dissolved in the crystallization buffer at 2.5 mg/mL. In the next step, 40 μL of the solution of the hirudin fragment was mixed with 160 μL of the thrombin solution. After incubation for 1 h at 4 °C, crystallization was carried out at 4 °C by the hanging-drop method. The hirudin/thrombin solution was mixed 1:1 with the reservoir solution (20 mM NaH2PO4, 27% polyethylene glycol 8000, pH 7.5) and 2 μL of this solution was placed to the center of a cover slide. Immediately after mixing of protein and reservoir buffer, microseeding was performed with a horse hair. The wells of the crystallization trays were filled with 500 μL of the reservoir buffer. Subsequently the coverslips were placed over the wells and sealed. Crystals of good diffraction quality could be obtained after 10–14 days.

For soaking, a DMSO stock solution of the inhibitors (A, B, and C (50 mM) and fragments 1–6 (500 mM)) was diluted 1:10 with a solution containing crystallization and reservoir buffer at a 1:1 ratio resulting in the final soaking concentration containing 5 or 50 mM of the respective compound and 10% DMSO. A crystal without visible imperfections was selected and transferred into the soaking solution for 24 h.

**Data Collection and Processing.** For data collection at 110 K, crystals were prepared using a cryoprotectant solution of 20% glycerol in reservoir buffer for 2 min.

The data sets of 1, 2, 5, and A were collected at the synchrotron DESY (Hamburg) at beamline P14, using a wavelength of 0.976 Å and a Pilatus 6M detector. Data sets of 3, 4, 5, and C were collected at the synchrotron BESSY (Berlin) at beamline 14.2 (wavelength of 0.9184 Å) and beamline 14.3 (wavelength of 0.895 Å), both using Rayonix MX-225 detectors.

For compounds 1, 2, 3, 5, 6, A, and C, data processing and scaling was performed with the XDS program package.27 For fragment 4, data processing and scaling was performed using the HKL2000 package.28

**Structure Determination and Refinement.** The coordinates of human thrombin (PDB codes 1H8D29 and 4UE7) were used for molecular replacement with Phaser from the CCP4 program package.30

For initial rigid-body refinement of the protein molecule, followed by repeated cycles of maximum-likelihood energy minimization simulated annealing and B-factor refinement, the program PHENIX31 was used. For determining the temperature factors for the structures anisotropic B-factor refinement was applied. A randomly chosen 5% set of all data was used for the calculation of Rmerge and not used during the refinement. Amino acid side chains were fit into σ-weighted 2Fo − Fc and Fo − Fc electron density maps using Coot.32 After the first refinement cycle, water molecules and subsequently ions and ligands were located in the electron density and added to the model. Restraints were applied to bond lengths and angles, planarity of aromatic rings, and van der Waals contacts. Multiple side chain conformations were built in case an appropriate electron density was observed and maintained during the refinement, and if the minor populated side chain showed at least 20% occupancy. The final models were validated using PHENIX validation options and Coot. The Ramachandran plot was calculated with MolProbity33 integrated in the PHENIX package. Data collection, unit cell parameters, and refinement statistics are given in Table 1. Analysis of temperature
factors was performed with MolProbity. The naming of the protein amino acids was done according to Bode et al.29

**Protein Data Bank and Accession Codes.** Coordinates and structure factors have been deposited in the Protein Data Bank (PDB) with the following accession codes: 1 (SAFP), 2 (4UD9), 3 (4UEH), 4 (4UE7), 5 (SAFY), 6 (SAHG), A (4UDW), and C (SAFZ).

**ASSOCIATED CONTENT**

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.5b00812.

Figure A, thermodynamic parameters determined by direct ITC titrations for the ligands A, B, and C in the presence and absence of Hirudin; Figure B, cluster dendrogram (PDF)

Formula strings (CSV)

**AUTHOR INFORMATION**

*Corresponding Author*

*Phone: +49-6421-28-21313. E-mail: klebe@mailer.unimarburg.de*

*Notes*

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

The authors thank Dr. Martina Schäfer and Marie Fricke (Bayer Healthcare, Berlin) for the initial discussion of the thermodynamic and structural data. Furthermore, we gratefully acknowledge the beamline staff at BESSY II (Helmholtz-Zentrum Berlin) in Berlin and Desy (EMBL) in Hamburg, Germany, for providing us outstanding support during the data collection. We also thank the Helmholtz-Zentrum Berlin for the travel support. The present study was kindly supported by the ERC Advanced Grant No. 268145-DrugProfIIbind awarded to G.K.

**ABBREVIATIONS USED**

FBLD, fragment-based lead discovery; ITC, isothermal titration calorimetry; S1, S2, S3, ..., P1, P2, P3, ..., classification of pockets and ligand positions according to the Berger and Schlechter notation used for proteases

**REFERENCES**


