J Mol Model (2007) 13:247–254
DOI 10.1007/s00894-006-0136-9

ORIGINAL PAPER

Thrombin inhibitors with novel P1 binding pocket functionality: free energy of binding analysis

Gregor Mlinsek · Marko Oblak · Milan Hodoscek · Tom Solmajer

Received: 3 April 2006 / Accepted: 12 July 2006 / Published online: 30 September 2006

Abstract The high incidence of thrombembolic diseases justifies the development of new antithrombotics. The search for a direct inhibitor has resulted in the synthesis of a considerable number of low molecular weight molecules that inhibit human α-thrombin potently. However, efforts to develop an orally active drug remain in progress as the most active inhibitors with a highly basic P1 moiety exhibit an unsatisfactory bioavailability profile. In our previous work we solved several X-ray structures of human α-thrombin in complexes with (1) novel bicyclic arginine mimetics attached to the glycylproline amide and pyridinone acetamide scaffold and (2) inhibitors with a novel aza scaffold and with charged or neutral P1 moieties. In the present contribution, we correlate the structures of the complex between these inhibitors and the protein with the calculated free energy of binding. The energy of solvation was calculated using the Poisson–Boltzmann approach. In particular, the requirements for successful recognition of an inhibitor at the protein's active site pocket S1 are discussed.

G. Mlinsek · M. Hodoscek · T. Solmajer
National Institute of Chemistry,
POB 660, Hajdrihova 19,
1001 Ljubljana, Slovenia

G. Mlinsek
Clinical Center Ljubljana,
Zaloska 2,
1000 Ljubljana, Slovenia

M. Oblak · T. Solmajer (✉)
Lek, d.d., Pharmaceutical and Chemical Company,
1526 Ljubljana, Slovenia
e-mail: tom.solmajer@ki.si

Keywords Thrombin · Inhibitor · Interaction energy · Poisson–Boltzmann · Factorising binding affinity · Epsilon estimate in protein interior

Introduction

The high incidence of thrombembolic diseases justifies the need for the development of new antithrombotics. The majority of existing human α-thrombin inhibitors are based on the natural substrate fibrinogen and belong to reversible di-and tripeptidic noncovalent inhibitors that are derived from the classical tripeptide D-Phe-Pro-Arg [1, 2]. However, in order to obtain an orally active drug, further modifications were made that resulted in very heterogeneous chemical structures. Two principle modifying strategies were used: (1) the highly basic P1 moiety was kept unchanged while the P2–P3 part was modified to increase bioavailability, (2) the basicity of the P1 part was lowered to improve bioavailability, but the loss in potency had to be compensated through modifications in the P2–P3 part [3]. The binding pockets of the thrombin active site are shown schematically in Fig. 1.

In order to introduce such novel chemical and biological properties, Kikelj et al. [4–6] prepared several new heterobicyclic P1 moieties of low basicity appended to the glycylprolyl scaffold [7] or pyridinone acetamide ring [8–10], which is devoid of *cis*/*trans* izomerization and thus conformationally more restricted. Zega et al. [11] incorporated an azapeptide moiety into the central part of the inhibitor structure. To improve the bioavailability, the basicity at the P1 site was reduced by exchanging benzamidine with benzamidoxime in some of these aza analogues.

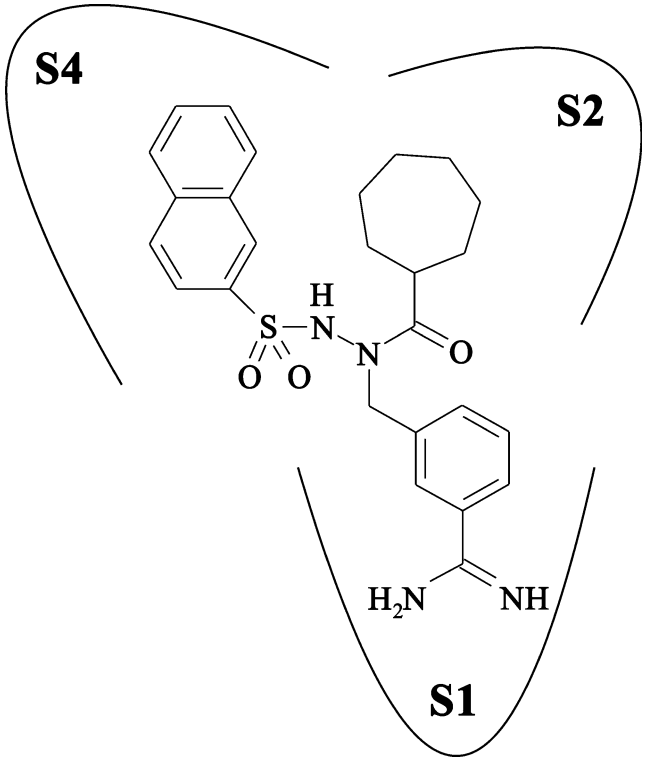


Fig. 1 Binding pockets of α-thrombin's active site

We have solved several X-ray structures of human α-thrombin in complexes with these novel reversible bicyclic arginine mimetics attached to the glycylproline amide and pyridinone acetamide scaffold (molecules 3 and 4 in Fig. 2) and inhibitors with the novel aza scaffold and neutral P1 moieties (molecules 1 and 2 in Fig. 2). The X-ray structures were disclosed and discussed [4–6, 11]. In the present paper we correlate the structures of ligands in the enzyme active site to their binding constants. Chemical groups or properties that were found to contribute to strong or weak binding of these inhibitors are pointed out/highlighted. An attempt is made to interpret the molecular recognition process through the calculated energy of interaction between the inhibitors and protein.

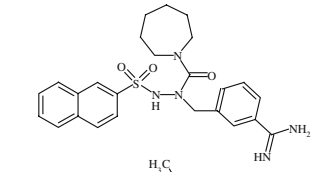
The correlation of binding affinity with the structural characteristics of a molecular system is an essential element of structure-based drug discovery. Estimation of the binding affinity from the structural parameters is not clear-cut and exact correlation between the structural and thermodynamic features has not yet been established [12–14]. In order to interpret our experimental data on α-thrombin–inhibitor complexes, obtained using protein crystallography, we have employed a physical model that involves a static description for interaction of molecules in the ligand–enzyme complex at the atomic level and a continuum solvent model [15, 16]. In principle, the use of X-ray or docked structures as the basis for parameterization of the interactions between ligands and proteins is in contrast to the frequently employed models of structures minimized by the molecular dynamics or Monte Carlo methods. We have opted to use the former model in order to provide a parallel/an analogue to parameterization of interactions in scoring function developments [17–19]. As electrostatic interactions are of crucial importance in understanding biochemical systems, and as the electrostatic complementarity between α-thrombin and its substrates or inhibitors seems crucial for successful recognition, we initially compared the total electrostatic contribution, including the Coulombic interaction and solvation contribution to the free energy of binding of the P1 moieties in a series of thrombin inhibitors in the S1 pocket of human α-thrombin. The former was expected to be energetically favorable while the latter is, in principle, an energetically costly process. The differences in a series of ligand–thrombin complex structures were analyzed, quantified and an explanation for the weaker binding constant for several of them was sought. In our model, the binding free energy is written according to the standard form introduced by Honig et al. [15, 16]. A more detailed description of the procedure is given in the Materials and methods section.

A significant problem in treating a system like this is the value of the dielectric constant for the protein interior, ε_i, which occurs in $\Delta G_{coul\ (\varepsilon i)}$ and $\Delta G_{solv\ (\varepsilon i,\ \varepsilon o)}$. Various approaches for estimating the dielectric constants at the protein's active site have been reported in the literature. It has been argued that values of 2–4 are the most appropriate for use in Poisson–Boltzmann calculations, because the contribution of ionizable groups to the fluctuations in dipole moment should not be included in the dielectric constant [20]. On the other hand, pK_a values computed assuming a value of 20 for the dielectric constant agree much better with the measured ones. We have chosen to estimate the dielectric constant value for α-thrombin's active site on the basis of structural data for α-thrombin–inhibitor complexes coupled to experimentally-determined thermodynamic data for these ligands when bound to α-thrombin's active site [21]. The parameterized dielectric constant was subsequently used for calculation of the individual energy contribution to the binding affinity of inhibitors 1–4 with structurally diverse P1 moieties.

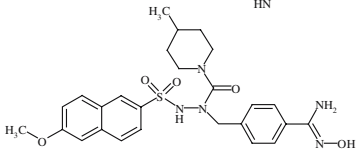
Materials and methods

The free energy of binding, ΔG_{bind}, of the system is given by Eqs. (1, 2, 3, 4, 5, 6, 7).

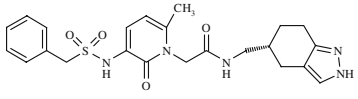
$$\Delta G_{bind} = \Delta H - T\Delta S \quad (1)$$



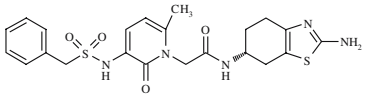
Molecule 1
Ki = 5 nM
Resolution: 1.73 Å
Reference: 11



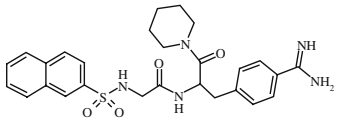
Molecule 2
Ki = 378 nM
Resolution: 2.7 Å
Reference: 11



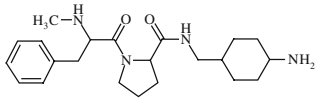
Molecule 3
Ki = 200 nM
Resolution: 1.9 Å
Reference: 6



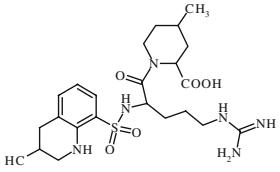
Molecule 4
Ki = 240 nM
Resolution: 2.0 Å
Reference: 4



NAPAP (1DWD.pdb)
Ki = 6 nM
Resolution: 3.0 Å
Reference: 12



L371912 (1TOM.pdb)
Ki = 5nM
Resolution: 1.8 Å
Reference: 13



Argatroban (1DWC.pdb)
Ki = 39 nM
Resolution: 3.0 Å
Reference: 12

Fig. 2 Chemical structures of α-thrombin inhibitors 1–4 [4–6, 11] and several other α-thrombin inhibitors published in the literature [34, 35]

$$\Delta G_{bind} = \Delta G_{el(\varepsilon i,\varepsilon o)} + \Delta G_{np} + \Delta G_{strain} - T\Delta S_{mc} + T\Delta S_{sc} + T\Delta S_{t,r} \quad (2)$$

$$\Delta G_{el(\varepsilon i,\varepsilon o)} + \Delta G_{coul(\varepsilon i)} + \Delta G_{solv(\varepsilon i,\varepsilon o)} \quad (3)$$

$\Delta G_{\mathrm{el}\ (\varepsilon i,\ \varepsilon o)}$ is the total change in the electrostatic free energy for two reactants with interior dielectric constant ε_i, both of which are embedded in a solvent of dielectric ε_o. It is obtained from the sum of the Coulombic and solvation energies.

$$\Delta G_{coul} = \sum_i \Delta G_{coul,i} = 1/2 \sum_i \sum_{j \neq i} q_i q_j / \varepsilon_i r_{ij} \quad (4)$$

$$\Delta G_{coul(\varepsilon i)} \approx E_{PI(\varepsilon I)} - E_{I(\varepsilon I)} - E_{P(\varepsilon I)} \quad (5)$$

$\Delta G_{\mathrm{coul}(\varepsilon i)}$ is the pairwise Coulombic free energy between two interacting molecules that are embedded in a medium of dielectric constant ε_i. The Coulombic energy term employed is shown in Eq. (4). It is calculated as the difference between the energy of the complex (PI) and that of the isolated reactants (protein (P), inhibitor (I)) (Eq. 5) [22]. Figure 3 shows the two cycles used for computing the binding free energy [23].

$\Delta G_{\mathrm{solv}\ (\varepsilon i,\ \varepsilon o)}$ is the change in the solvation free energy upon transferring the various species from a medium of dielectric ε_o=1 (vacuum) to a medium of dielectric ε_o=78.5 (water). It is calculated from thermodynamic cycle 1, which is shown in Fig. 3. The electrostatic contribution to the solvation binding energy was obtained using the Poisson–Boltzmann approach as implemented in CHARMm [24]. Energy values, $E_{i,j}$, were calculated for each of the six states listed in Fig. 3. The solvation energy values of the inhibitor, protein and the complex were calculated in vacuum and in water. The values obtained in water were subtracted from the values obtained in vacuum to give the solvation energies for the inhibitor, protein and the complex, respectively, as shown in Eq. (6). (E=energy, v=vacuum, w=water).

$$\Delta G_{solv(\varepsilon i,\varepsilon o)} \approx \left(E_{PI,v} - E_{PI,w}\right) - \left(E_{I,v} - E_{I,w}\right) - \left(E_{P,v} - E_{P,w}\right) \quad (6)$$

The nonpolar (hydrophobic) contribution to the binding free energy is approximated by Eq. (7), where γ_{aw} is the microscopic surface tension associated with the transfer of an alkane from a liquid alkane to water. To calculate this nonpolar contribution to the solvation free energy, van der Waals interactions are not considered explicitly but are assumed to be taken into account implicitly by the surface tension parameter, γ_{aw}.

$$\Delta G_{np} \approx \gamma_{aw}\Delta(CA) = \left(\gamma_{aw}A_{C,w}\right) - \left(\gamma_{aw}A_{I,w}\right) - \left(\gamma_{aw}A_{P,w}\right) \quad (7)$$

The curvature correction factor, C, describes the effect of the curvature of the molecule on its interfacial free energy [25], while A is the solvent accessible surface area [26]. Several γ_{aw} values (0.03–0.06 kcal mol^{-1} Å^{-2}), depending on the force field employed in the calculation, have been reported in the literature [23, 25]. In principle, the γ_{aw} value depends strongly on the radii used to calculate the accessible surface area. Recently, Im et al. [27] used the all-hydrogen PARM22 potential function of CHARMm and a surface tension coefficient 0.03 kcal mol^{-1} Å^{-2}. In our experiment we used the same potential function and parameter set and thus the same γ_{aw} value.

$\Delta G_{\mathrm{strain}}$ (Eq. 2) accounts for possible distortions in the system upon complexation. $T\Delta S_{\mathrm{mc}}$ and $T\Delta S_{\mathrm{sc}}$ describe the

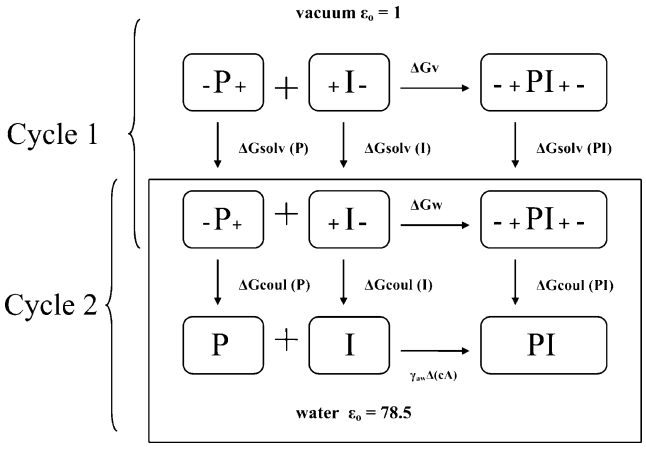


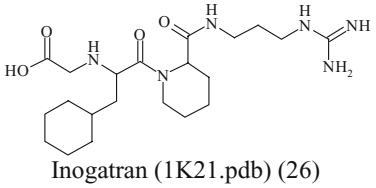
Fig. 3 Thermodynamic cycles for calculating the binding free energy of the protein–inhibitor complex formation. ΔG_{solv} is the electrostatic contribution to the solvation free energy of the protein (P), inhibitor (I) and the complex (PI). It is calculated as shown in cycle 1. ΔG_v and ΔG_w are the free energies of complex formation in vacuum and in water

loss of configurational entropy due to loss of backbone and side-chain torsional freedom and $T\Delta S_{t,r}$ accounts for the loss of translational and rotational degrees of freedom upon binding [15, 16, 22].

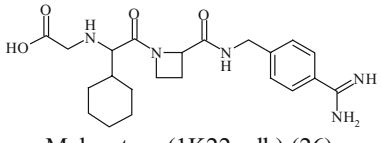
All molecular mechanics calculations presented in this work were performed using the CHARMM program package [24] and the CHARMm22 force field. For each of the molecules from the PDB source used in this work the hydrogens of the low molecular weight inhibitors were added manually and initially positioned by HBUILD and then atom types for the inhibitor were assigned. The structure was energy minimized by 200 steepest descents (SD) steps and subsequently by 800 Newton Ralphson (ABNR) steps, with only the hydrogens allowed to move. For the structure thus obtained, in vacuo Mulliken single point charges were calculated by GAMESS [28] using the standard 6-31G(d) basis set. Each ternary complex α-thrombin–hirugen–inhibitor was then soaked in water to produce a 6.5 Å thick layer around the surface. The complex was energy minimized by 200 SD and 800 ABNR steps, with only the hydrogens and the molecules added in the water coat allowed to move and in vacuo charges used for the inhibitor. A residue-based cutoff of 12 Å was used in the MM calculation. In the next step, Mulliken single point charges for the inhibitor were calculated de novo using the QM/MM methodology [29] in order to include the influence of the surrounding protein. The inhibitor was treated by quantum and the protein by molecular mechanical means. The basis set for QM/MM calculations was 6–31G(d). With the new charges, we performed a third cycle of energy minimization using the protocol described above. This cycle was employed to minimize hydrogens taking into account the optimized charges of the inhibitor.

Estimation of the epsilon value

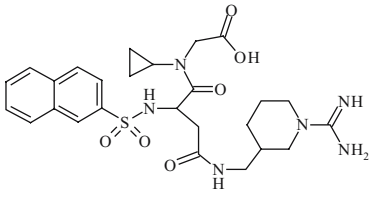
To find the molecular system most appropriate for parameterization of the dielectric constant we used three X-ray structures (Fig. 4) whose enthalpies of binding have been published in the literature [21] and which have similar structures to our set of molecules 1–4. Three X-ray structures were considered for this purpose i.e. inogatran (1K21.pdb), melagatran (1K22.pdb) and napsagatran, which was kindly provided by Dr. D. Banner (Roche). Experiments were performed on the structures obtained after 100 steps of SD minimization of all atoms, to relax the strains while keeping the geometry of the inhibitor almost unchanged (rmsd for the complex is 0.1–0.2 Å). Entropy can significantly affect the value of the free energy of binding. It was shown experimentally that an entropic contribution can oppose an enthalpic contribution (napsagatran, inogatran) or can add 20% to the binding affinity (melagatran) or up to 40% in an analogue of napsagatran [21]. However, the calculation of its exact value is notoriously difficult. Thus, to estimate the dielectric constant value from the X-ray structures and thermodynamic data we have used the experimental ΔH values [21]. The most problematic expression, ΔG_{np}, which is a mixed enthalpic and entropic term, was either included or omitted. Its incorporation gave an overestimated dielectric constant value and thus its upper boundary while its omission results in an underestimated value. We hypothesized that by using both approaches we would get the best estimate for this parameter. The value of the dielectric constant was systematically varied until relation (8) was satisfied (Table 1). Relation (8) is an approximation based on the



Inogatran (1K21.pdb) (26)



Melagatran (1K22.pdb) (26)



Napsagatran (not deposited)

Fig. 4 Molecules used for parameterization of the ε value

assumption that $\Delta G_{coul(\varepsilon i)}$ and $\Delta G_{solv(\varepsilon i,\ \varepsilon o)}$ represent energetic and not entropic contributions to ΔG_{bind}. The sum of $\Delta G_{coul(\varepsilon i)}$ and $\Delta G_{solv(\varepsilon i,\ \varepsilon o)}$ is thus correlated with the experimental value of ΔH_{bind} and not ΔG_{bind}. For further parameterization we have therefore used an average value of 2.7.

$$\Delta Hbind_{exp} \approx \Delta G_{coul(\varepsilon i)} + \Delta G_{solv(\varepsilon i,\varepsilon o)} \left(+\Delta G_{np}\right) \quad (8)$$

The method that we have used to factorize the binding affinity is parameter-dependent. During the Poisson–Boltzmann calculation we used CHARMm's parameter set PARM22, a grid spacing of 0.4 Å, a salt concentration of 0.15 mM and a γ value of 0.03 kcal mol^{-1} Å^{-2}. A grid with a unit dimension of 0.4 Å was used to limit the CPU time needed. The convergence of the grid unit was estimated using a grid of 0.25 Å and no significant changes were observed in the computed quantities. Further tests with other atomic radii [30] gave no improvement in the results.

Results and discussion

The experimental observation of similarity in chemical structures with a similar binding mode but with a considerable difference in the binding constant calls for an explanation of the weakness of some of the inhibitors and the strength of the others.

Reversible α-thrombin inhibitors are kept in the active site through long range i.e. electrostatic and van der Waals interactions. Ionic hydrogen bonds are best characterized since the ionic contacts between Asp189 and the P1 moiety of the inhibitor are of crucial importance for effective binding of α-thrombin inhibitors. Ionic contacts in parallel with the number and strength of the hydrogen bonds formed between the functional groups of the ligand and protein are believed to give at least a qualitative correlation between the structure and the potency of the inhibitor. However, certain inhibitors can bring about very strong binding constants, although they form only a few hydrogen bonds and no salt bridge to α-thrombin [31]. The reasons for the efficacy of such binding can be low solvation cost or important weak nonbonding interactions. Unique inhibition motives, such as a cluster of very short hydrogen bonds (<2.3 Å) in the catalytic triad, have also been proposed [32]. Furthermore, the entropic contribution to the free energy of binding can also be of significance. It can oppose the enthalpic contribution or comprise up to 40% to the free energy of binding [33]. A clear-cut correlation between structural or geometric features and changes in binding enthalpy or entropy is thus not straightforward.

We have directed our efforts to the understanding of molecular recognition that is based on a structural description by elucidating the recognition process in terms of the energy of binding. The relative contribution of the fragments of individual inhibitors to the energy of binding in a homologous series was computed to enable us to decipher the reasons for weak or strong binding. The method that we used to factorize the binding affinity is parameter-dependent and sensitive to the coordinates of the complex. A minimization that brings about RMS changes of only 0.1–0.2 Å can cause considerable changes in the value of the Coulombic term. Even the use of X-ray structures with no minimization of heavy atoms reduced this problem only to a certain degree, bearing in mind that resolution of the structures solved was not equal in all cases. In those cases where the complex is treated as charged it is easier to draw conclusions as the values of the total electrostatic component are considerably above/below zero compared to uncharged complexes where the electrostatic energy values are close to zero and are therefore also sign dependent. Resultant values should therefore be regarded with caution. Exact values are thus of less importance than trends that can be derived from the calculated values.

Inspection of the X-ray structures shows that the geometry of inhibitors **1**–**4** enables their proper binding to the active site [4–6, 11, 34]. They form hydrogen bonds that are frequently observed in human α-thrombin–inhibitor complexes. While molecule **1** (Ki=5 nM) has a low nanomolar potency, Ki values of **2** (Ki=378 nM), **3** (Ki=200 nM) and **4** (Ki=240 nM) are in the high nanomolar range. We conjectured that the reason for this could be low basicity of the inhibitor's P1 moiety which results in nonionic hydrogen bonds in the thrombin active site. In order to obtain enhanced insight into the relative importance of individual energy terms we have factorized the binding energy (Eq. 2). Our major interest was in the electrostatic component that is believed to be of crucial importance for α-thrombin inhibition. In order to calculate the total electrostatic component composed of the Coulombic interaction and the solvation term, an appropriate value for the protein's dielectric constant was required. As described in the Materials and methods section, we parameterized the dielectric constant through use of the experimental X-ray structures of α-thrombin in complex with inhibitors of similar structure: napsagatran ($\Delta H bind_{exp}$ =

Table 1 Parameterization of α-thrombin protein dielectric constant (the influence of the nonpolar term ΔG_{np} is shown)

ΔG_{np}	Napsagatran ε_i	Inogatran ε_i	Melagatran ε_i	Average ε_i
Excluded	1.9	2.3	2.5	2.2
Included	2.5	3.1	3.6	3.1

−70 kJ mol^{-1}), melagatran ($\Delta H bind_{exp}$ =−57.8 kJ mol^{-1}) and inogatran ($\Delta H bind_{exp}$ =−37.0 kJ mol^{-1}) (Fig. 4). The purpose of this parameterization is estimation of the ε value which enables further ΔG calculations. The results of this calculation are presented in Table 1.

All calculated values fall within the range of the calculated dielectric constants on folded proteins, which were reported to be between 2 and 5 [20].

If ΔG_{np} is omitted, the ε value is underestimated due to neglect of the enthalpic part of the nonpolar term. On the contrary, the value of ε obtained by inclusion of the ΔG_{np} term is overestimated.

For further calculation, we choose an average value of 2.7. As shown in Table 2, calculations using a theoretical dielectric constant value of 2 result in the same trends as calculations obtained with a value of 2.7, further accentuating the importance of electrostatics. An ε value of 2 is a theoretical value for the case where the effects of electronic polarizability are included and close to a value of 1 to which the CHARMm force field is parameterized.

Table 2 Coulombic interaction (ΔG_{Coul}), electrostatic contribution to the solvation free energy of binding (ΔG_{solv}), their sum (ΔG_{el}), nonpolar contribution to the solvation free energy (ΔG_{np}) and the total sum of these terms ($\Delta G_{Coul}+\Delta G_{solv}+\Delta G_{np}$) for α-thrombin inhibitors and their P1 moieties only; dielectric constants of 2.7 and 2 were used (N — number of atoms in the ligand)

Molecule	ε_i	System	N	ΔG_{Coul}	ΔG_{solv}	ΔG_{el}	ΔG_{np}	Total
Argatroban 	2.7	Inhibitor-protein	71	-36.54	29.04	-7.50	-9.55	-17.05
	2.7	P_1-protein	18	-32.16	19.78	-12.38	-2.66	-15.04
	2.7	P_1-aspartate	18	-37.92	17.42	-20.50	-0.41	-20.91
	2	Inhibitor-protein	71	-49.33	31.08	-18.25	-9.55	-27.80
L371912 	2.7	Inhibitor-protein	63	-45.38	38.87	-6.51	-10.20	-16.71
	2.7	P_1-protein	20	-38.20	26.12	-12.08	-3.65	-15.73
	2.7	P_1-aspartate	20	-34.81	12.04	-22.77	-0.19	-22.96
	2	Inhibitor-protein	63	-61.27	42.08	-19.19	-10.20	-29.39
NAPAP 	2.7	Inhibitor-protein	69	-54.38	39.85	-14.53	-10.88	-25.41
	2.7	P_1-protein	17	-44.74	29.27	-15.47	-3.60	-19.07
	2.7	P_1-aspartate	17	-46.30	25.43	-20.87	-0.61	-21.48
	2	Inhibitor-protein	69	-73.42	42.81	-30.61	-10.88	-41.49
1 	2.7	Inhibitor-protein	64	-55.61	41.29	-14.32	-10.77	-25.09
	2.7	P_1-protein	17	-48.33	29.40	-18.93	-4.51	-23.44
	2.7	P_1-aspartate	17	-44.04	25.06	-18.98	-0.58	-19.56
	2	Inhibitor-protein	64	-75.07	45.14	-29.93	-10.77	-40.70
3(+) 	2.7	Inhibitor-protein	61	-54.74	40.36	-14.38	-12.23	-26.61
	2.7	P_1-protein	19	-45.59	23.85	-21.74	-4.53	-26.27
	2.7	P_1-aspartate	19	-44.62	22.18	-22.44	-0.63	-23.07
	2	Inhibitor-protein	61	-73.91	43.40	-30.51	-12.23	-42.74
3 	2.7	Inhibitor-protein	60	-16.47	23.72	7.25	-12.24	-4.99
	2.7	P_1-protein	18	-7.38	8.53	1.15	-4.54	-3.39
	2.7	P_1-aspartate	18	-9.17	7.31	-1.86	-0.63	-2.49
	2	Inhibitor-protein	60	-22.23	25.90	3.67	-12.24	-8.57
2 	2.7	Inhibitor-protein	68	-14.57	24.13	9.56	-12.61	-3.05
	2.7	P_1-protein	17	-13.45	15.33	1.88	-5.31	-3.43
	2.7	P_1-aspartate	17	-8.55	9.03	0.48	-0.56	-0.08
	2	Inhibitor-protein	68	-19.67	25.74	6.07	-12.61	-6.54
4(+) 	2.7	Inhibitor-protein	59	-49.45	37.98	-11.47	-12.55	-24.02
	2.7	P_1-protein	20	-46.47	23.62	-22.85	-4.95	-27.80
	2.7	P_1-aspartate	20	-33.98	10.95	-23.03	-0.28	-23.31
	2	Inhibitor-protein	59	-66.76	41.10	-25.66	-12.55	-38.21
4 	2.7	Inhibitor-protein	58	-12.38	25.62	13.24	-12.55	0.69
	2.7	P_1-protein	19	-12.29	12.73	0.44	-4.95	-4.51
	2.7	P_1-aspartate	19	-5.18	3.99	-1.19	-0.28	-1.47
	2	Inhibitor-protein	58	-16.72	27.62	10.90	-12.55	-1.65

Table 2 shows the Coulombic energy and electrostatic contributions to the solvation free energy, their sum, nonpolar contribution to the solvation free energy and the overall sum of all these terms for the following interaction systems: protein-inhibitor, P1 moiety-protein and P1 moiety-Asp189 of various inhibitors. The latter was considered as an isolated system. The dielectric constants used were 2.7 and 2.0. The P1 moiety is treated as depicted in the structure in Table 2. All calculations were performed on those X-ray structures where only the added hydrogen atoms were allowed to move during a short minimization procedure. The results in Table 2 show that all Coulombic terms and all nonpolar solvation free energy terms are negative in sign and all electrostatic solvation free energy terms are positive regardless of whether the molecule was treated as charged or uncharged. However, the sum of the Coulombic and electrostatic solvation terms differ with the charge status of the molecule and with the system (inhibitor-protein or P1 moiety-inhibitor) which shows the different importance of the Coulombic and solvation term in these cases.

The interaction of the highly basic P1 moiety with the protein treated in the charged form is favorable. The same trend as in the P1 moiety-protein system is also seen in the isolated P1 moiety-Asp189 system, which is in agreement with the importance of Asp189 as one of the crucial residues in this molecular system and the origin of this favorable electrostatic contact. Molecules 3 and 4, when treated as charged, also show high negative values in both the P1-protein and the P1-Asp189 systems. However, the same molecules, if treated as uncharged, result in slightly positive values in the P1 moiety-protein and P1 moiety-Asp189 systems, indicating that in this case the electrostatic component of binding is unfavorable. The sum of the Coulombic and electrostatic solvation energies of the whole inhibitor is negative in argatroban, L317912, NAPAP, 1, 3 and 4 if treated as charged. Low pK_a values of the P1 moiety of 3 and 4 tell us that the uncharged molecular form prevails at pH 7.4 in these two cases. If molecules 2–4 are treated as uncharged, the overall electrostatic interaction is clearly positive. In addition, the value of the P1 moiety-protein is slightly above zero. In this case the absolute value of both energy terms – Coulombic and electrostatic – falls, with the former term falling relatively more. The Coulombic term is still negative but too small to overcome the cost of solvation. The same is seen with 2, which only exists uncharged and decays to benzamidine on protonation. However, together with the ΔG_{np} term, these P1 moiety-protein and inhibitor-protein systems show slightly negative terms overall, which is in agreement with the lower potency of these inhibitors and with the view that regarding such inhibitors, parts other than P1 are crucial for effective binding. As P2 and P3 of 3 and 4 are identical and also occupy identical positions in the X-ray structure, and as both P1 moiety-proteins show the same sign and size the slightly positive value of 0.69 for the inhibitor-protein in 4 can be explained as an error.

The nonpolar solvation free energy has almost the same value in the P1 moiety-protein system regardless of the charge status of the inhibitor. However, in the uncharged form it is of higher importance as it turns the total sum of the three energy terms negative and is thus favorable, while in charged molecules it adds only slightly to the highly negative sum of the Coulombic and solvation terms. As the contact between nonpolar parts in the P1 moiety-Asp189 system is negligible, the calculated value is also as such.

The calculated values thus show trends that are in agreement with the experimental data, i.e. the Ki values of the molecules shown in Fig. 2. However, no generalization about the binding constant value can be made based on these results, as several entropy terms from Eq. (2) were not considered.

Conclusions

Due to the lower basicity at the heterobicyclic P1 moiety, molecules 1–4 form weaker electrostatic contacts to Asp189. Although the solvation cost of the uncharged P1 moieties decreases, the Coulombic interaction increases relatively more so that the overall electrostatic contribution for the P1 moiety-protein and inhibitor-protein turns positive and is thus unfavorable. Our results show, in semiquantitative terms, that in molecules with an uncharged P1 residue, complex formation with α-thrombin is lead/guided by forces other than electrostatics, which is in agreement with the conclusion that in order to increase the binding of such inhibitors the interactions of the hydrophobic P2 and P3 moieties need to be optimized.

Acknowledgements We would like to thank Dr. D. Banner (Roche) who kindly provided the coordinates of the napsagatran-α-thrombin complex and Dr. J. Mavri (NIC, Ljubljana) for helpful comments and discussions.

References

1. Vacca JP (1998) Annu Rep Med Chem 81–90
2. Fevig JM, Wexler RR (1999) Annu Rep Med Chem 81–100
3. Rewinkel JBM, Adang AEP (1999) Curr Pharm Des 5:1043–1075
4. Marinko P, Krbavcic A, Mlinsek G, Solmajer T, Bakija AT, Stegnar M, Stojan J, Kikelj D (2004) Eur J Med Chem 39:257–265
5. Peterlin-Masic L, Mlinsek G, Solmajer T, Trampus-Bakija A, Stegnar M, Kikelj D (2003) Bioorg Med Chem Lett 13:789–794
6. Peterlin-Masic L, Kranjc A, Mlinsek G, Solmajer T, Stegnar M, Kikelj D (2003) Bioorg Med Chem Lett 13:3171–3179
7. Tucker TJ, Lumma WC, Mulichak AM, Chen ZG, Naylor–Olsen AM, Lewis SD, Lucas R, Freidinger RM, Kuo LC (1997) J Med Chem 40:830–832

8. Sanderson PEJ, Dyer DL, Naylor-Olsen AM, Vacca JP, Gardell SJ, Lewis SD, Lucas BJ, Lyle EA, Lynch JJ, Mulichak AM (1997) Bioorg Med Chem Lett 7:1497–1500
9. Sanderson PE, Cutrona KJ, Dorsey BD, Dyer DL, McDonough CM, Naylor-Olsen AM, Chen IW, Chen ZG, Cook JJ, Gardell SJ, Krueger JA, Lewis SD, Lin JH, Lucas BJ, Lyle EA, Lynch JJ, Stranieri MT, Vastag K, Shafer JA, Vacca JP (1998) Bioorg Med Chem Lett 8:817–822
10. Sanderson PE, Lyle TA, Cutrona KJ, Dyer DL, Dorsey BD, McDonough CM, Naylor-Olsen AM, Chen IW, Chen ZG, Cook JJ, Cooper CM, Gardell SJ, Hare TR, Krueger JA, Lewis SD, Lin JH, Lucas BJ, Lyle EA, Lynch JJ, Stranieri MT, Vastag K, Yan YW, Shafer JA, Vacca JP (1998) J Med Chem 41:4466–4474
11. Zega A, Mlinsek G, Solmajer T, Trampus-Bakija A, Stegnar M, Urleb U (2004) Bioorg Med Chem Lett 14:1563–1567
12. Solmajer T, Zupan J (2004) Drug discovery today, Technologies, 1. Lead optimization, pp 247–252
13. Mlinsek G, Novic M, Hodoscek M, Solmajer T (2001) J Chem Inf Comput Sci 41:1286–1294
14. Mlinsek G, Novic M, Kotnik M, Solmajer T (2004) J Chem Inf Comput Sci 44:1872–1882
15. Froloff N, Windemuth A, Honig B (1997) Prot Sci 6:1293–1301
16. Polticelli F, Ascenzi P, Bolognesi M, Honig B (1999) Protein Sci 8:2621–2629
17. Searle MS, Williams DH, Gerhard U (1992) J Am Chem Soc 114:10697–10704
18. Williams DH, Cox JPL, Doig AJ, Gardner M, Gerhard U, Kaye PT, Lal AR, Nicholls IA, Salter CJ, Mitchell RC (1991) J Am Chem Soc 113:7020–7030
19. Rarey M, Kramer B, Lengauer T, Klebe G (1996) J Mol Biol 261:470–489
20. Antosiewicz J, McCammon JA, Gilson MK (1996) Biochemistry 35:7819–7833
21. Dullweber F, Stubbs MT, Musil D, Sturzebecher J, Klebe G (2001) J Mol Biol 313:593–614
22. Gilson MK, Honig B (1988) Proteins 4:7–18
23. Friedman RA, Honig B (1995) Biophys J 69:1528–1535
24. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan, Karplus M (1983) J Comp Chem 4:187–217
25. Nicholls A, Sharp KA, Honig B (1991) Proteins 11:281–296
26. Lee B, Richards FM (1971) J Mol Biol 55:379–400
27. Im W, Beglov D, Roux B (1998) Comput Phys Commun 111:59–75
28. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery Jr JA (1993) J Comp Chem 14:1347–1363
29. Lee YS, Hodoscek M, Brooks BR, Kador PF (1998) Biophys Chem 70:203–216
30. Nina M, Beglov D, Roux B (1997) J Phys Chem B 101:5239–5248
31. Bone R, Lu T, Illig CR, Soll RM, Spurlino JC (1998) J Med Chem 41:2068–2075
32. Katz BA, Elrod K, Luong C, Rice MJ, Mackman RL, Sprengeler PA, Spencer J, Hataye J, Janc J, Link J, Litvak J, Rai R, Rice K, Sideris S, Verner E, Young W (2001) J Mol Biol 307:1451–1486
33. Zega A, Mlinsek G, Sepic P, Grdadolnik SG, Solmajer T, Tschopp TB, Steiner B, Kikelj D, Urleb U (2001) Bioorg Med Chem 9:2745–2756
34. Banner DW, Hadvary P (1991) J Biol Chem 266:20085–20093
35. Lyle TA, Chen ZG, Appleby SD, Freidinger RM, Gardell SJ, Lewis SD, Li Y, Lyle EA, Lynch JJ, Mulichak AM, Ng AS, NaylorOlsen AM, Sanders WM (1997) Bioorg Med Chem Lett 7:67–72