doi:10.1016/j.jmb.2007.01.022 J. Mol. Biol. (2007) 367, 1347–1356

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Protonation Changes upon Ligand Binding to Trypsin and Thrombin: Structural Interpretation Based on pK_a Calculations and ITC Experiments

Paul Czodrowski, Christoph A. Sotriffer and Gerhard Klebe*

Department of Pharmaceutical Chemistry, Philipps-University Marburg, Marbacher Weg 6 35032 Marburg, Germany

The protonation states of a protein and a ligand can be altered upon complex formation. Such changes can be detected experimentally by isothermal titration calorimetry (ITC). For a series of ligands binding to the serine proteases trypsin and thrombin, we previously performed an extensive ITC and crystallographic study and were able to identify protonation changes for four complexes. However, since ITC measures only the overall proton exchange, it does not provide structural insights into the functional groups involved in the proton transfer. Using Poisson–Boltzmann calculations based on our recently developed PEOE_PB charges, we compute pK_a values for all complexes of our former study in order to reveal the residues with altered protonation states. The results indicate that His57, a member of the catalytic triad, is responsible for the most relevant pK_a shifts leading to the experimentally detected protonation changes. This finding is in contrast to our previous assumption that the observed protonation changes occur at the carboxylic group of the ligands. The newly detected proton acceptor is used for a revised factorization of the ITC data, which is necessary whenever the protonation inventory changes upon complexation. The pK_a values of complexes showing no protonation change in the ITC experiment are reliably predicted in most cases, whereas predictions of strongly coupled systems remain problematic.

Keywords: protonation states; pK_a values; serine proteases; Poisson–Boltzmann; ITC

*Corresponding author

Introduction

The protein-ligand binding process is governed by several types of interactions, such as hydrogen bonds, charge interactions, aromatic stacking, and hydrophobic complementarity. What are often neglected for the sake of simplicity are protonation changes of the ligand and/or protein upon complex formation. Systematic studies considering such protonation effects have rarely been performed due to their experimental complexity. Isothermal titration calorimetry (ITC) represents one experimental method for the detection of overall changes in protonation. However, ITC reveals only the net stoichiometry for the exchange of protons and does not provide structural evidence about the groups that release or pick up the transferred protons. More detailed insights at the molecular level can be gained computationally by calculating pK_a values of functional groups and their changes upon complex formation. Calculations based on the Poisson–Boltzmann (PB) equation have achieved convincing agreement with experimental pK_a values observed for protein residues.[1–4] Other competitive approaches, such as the protein-dipole Langevin-dipole method (PDLD),[5–8] empirical models based on rules for hydrogen bonding contributions and solvation effects,[9,10] or simplified electrostatic models using sigmoidally screened Coulomb potentials,[11] have shown convincing agreement with experimental data. We recently developed a generally applicable charge model for PB-based pK_a calculations of protein–ligand complexes.[12] Here, we describe the application of this method to trypsin (Try) and thrombin (Thr) complexes (Figure 1),

Abbreviations used: ITC, isothermal titration calorimetry; PB, Poisson–Boltzmann; PDLD, protein-dipole Langevin-dipole; Try, trypsin; Thr, thrombin.

E-mail address of the corresponding author: klebe@mailer.uni-marburg.de

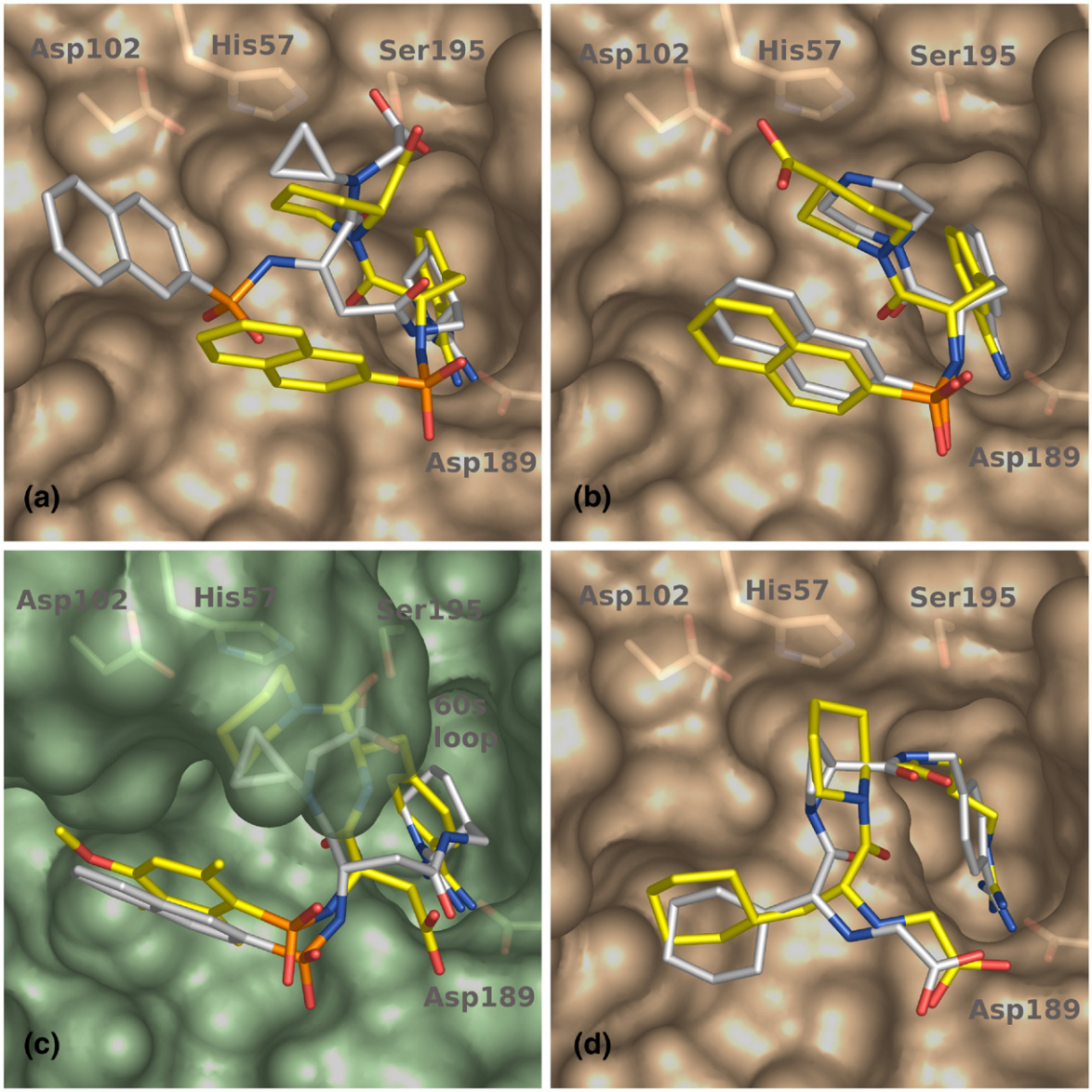

Figure 1. Trypsin (pale yellow) and thrombin (green) in complex with the following ligands: (a) **1b** (yellow) and **2** (white), (b) **1c** (yellow) and **1d** (white), (c) **2** (white) and **3** (yellow), (d) **4** (yellow) and **5** (white). The catalytic triad and Asp189 of the S1 pocket are shown in stick representation. The Figure was generated with PyMOL [http://pymol.sourceforge.net/].

which have been studied by ITC in our laboratory.[13] Evidence from the computational evaluations is used to interpret and factorize experimental microcalorimetric data recorded for a series of structurally related complexes.

Trypsin and thrombin belong to the family of trypsin-like serine proteases. They are well-established model systems for studying fundamental concepts of protein–ligand interactions. Both enzymes catalyze the peptide bond cleavage by means of their active-site residues His57, Asp102, and Ser195, which form the catalytic triad. Although the catalyzed reaction is identical, trypsin and thrombin belong to different biological pathways: Trypsin is involved in digestion, whereas thrombin is engaged in the blood-clotting cascade. Thrombin has been a target of pharmaceutical relevance for a long time, since its inhibition seems promising for the treatment of thrombosis *via* orally administered drugs. However, such drugs should be selective for thrombin and avoid undesired cross-reactivity with the structurally very similar trypsin. Therefore, the two proteins represent an ideal system to investigate the features determining selectivity. In Figure 2, the key characteristics of the binding pocket of trypsin/thrombin are shown: The S_1 pocket can accommodate a positively charged functional group, usually forming a salt-bridge to Asp189. Aromatic interactions are possible with the residues flanking the S_3 pocket. In contrast to trypsin, the extended 60-loop is present in thrombin and shields the S_2 pocket from the surrounding solvent environment. In both enzymes, the oxy-anion hole is found adjacent to the catalytic triad.

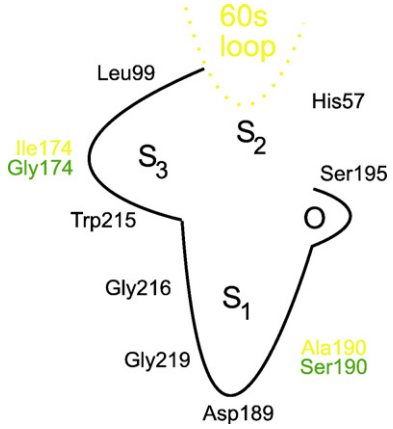

Figure 2. A representation of the trypsin/thrombin active site. The following color code is used for the origin of the amino acid residues: trypsin, green; thrombin, yellow; trypsin and thrombin, grey. The 60s loop is only present in thrombin.

A prerequisite for endowing ligands with enhanced target selectivity by rational concepts is the experimental factorization of affinity into entropic and enthalpic contributions, as obtained, for example, by means of ITC. Indirectly, this method also provides the overall change in protonation when a ligand binds to a protein. Here, we show results from a computational study that help us to understand and interpret unexpected protonation effects upon ligand binding to trypsin and thrombin. Such results are essential for correctly factorizing ITC data, because the heat of ionization required for this factorization depends on the type of titratable group. A meaningful partitioning into ΔH and $T\Delta S$ can be performed only if this heat effect is appropriately accounted for.

Earlier, we described the experimental characterization of the thermodynamic and structural properties of a series of trypsin and thrombin complexes.[13] For one thrombin and three trypsin complexes, a change of protonation has been detected during the ligand-binding process (Table 1). However, ITC and X-ray crystallography did not allow unequivocal assignment of the proton donor or acceptor. This assignment, which is crucial for the thermodynamic interpretation, is the main subject here.

Results

The results of the pK_a calculations will be discussed in light of the experimentally determined protonation changes Δn_{exp} (a positive sign corresponds to a proton uptake from the buffer upon complexation, whereas a negative sign symbolizes a release of protons; the value is given in units of mol protons/mol titratable group). Only pK_a shifts between the apo and the complexed form that result in a protonation change will be described. The corresponding degrees of protonation and the resulting change in protonation Δn_{calc} upon ligand binding is portrayed on the basis of the Henderson–Hasselbalch equation at a pH of 7.8, as used in all ITC experiments. Besides the calculated overall protonation change Δn_{calc}, the protonation change $\Delta n_{calc/HL}$ (based solely on pK_a shifts of His57 and the ligand) are evaluated. All results are summarized in Table 2. Further details about the pK_a calculations are available as Supplementary Data.

Estimating the accuracy of calculated protonation changes

Before the description of any individual results, we point out some general thoughts about protonation states and their possible changes, in light of an attempt to assess the significance and accuracy of calculated protonation changes.

In the context of the present study, we aim at the prediction of protonation changes occurring upon protein–ligand binding. Using a model scenario, we try to clarify what is a significant pK_a shift with respect to protonation changes. As an example, consider a titratable group with a pK_a of 7.0. In a buffer solution of pH 7.0 (i.e. the same value as the pK_a), the group is 50% protonated. If a pK_a shift of 1 unit is induced for this group (e.g. by the ligand-binding process), the protonation changes by ±41%, i.e. to 91% protonation if a shift of +1 pK_a unit is induced or to 9% protonation if the pK_a is shifted by −1 unit. This essentially results in a fully protonated or deprotonated group, respectively. In contrast, if the same shift of 1 unit is taking place at a group with a pK_a of 5.0, the effect is much less pronounced and virtually irrelevant for the protonation inventory (e.g. for pK_a 5.0→6.0 at pH 7.0, the protonation degree rises from 1% to 9%; for pK_a 5.0→4.0 at pH 7.0, it decreases from 1% to virtually zero. Thus, the maximum effect on the protonation inventory is exerted by groups with pK_a values near the pH.

Furthermore, the prediction accuracy of the computed pK_a values has to be considered. A pK_a shift of 1 unit is close to the estimated minimum deviation of predicted pK_a values from experimental data based on our calculation method (a root-mean-square deviation of 0.88 pK_a units was obtained predicting 132 experimentally determined pK_a values).[12] Therefore, with respect to estimates of protonation changes, caution is warranted if pK_a shifts of less than 1 unit are computed near the buffer pH. This means also that protonation changes of less than 40% (or 0.4 mol of transferred protons)

Table 1. Net change of protonation upon ligand binding to trypsin (Try) or thrombin (Thr)[13] as measured by ITC

Complex	Δn_{exp}
1b.Try	+0.90±0.06
1d.Try	−0.53±0.04
2.Try	+0.93±0.12
2.Thr	+0.88±0.08

A positive sign corresponds to a proton uptake from the buffer upon complexation, whereas a negative sign symbolizes a release of protons.

Table 2. Calculated protonation changes (Δn) along with the corresponding protonation degrees at pH 7.8 of the protein-ligand system before (apo) and after complexation (complexed)

					Protonation degree (pK_a) apo + ligand in water			Protonation degree (pK_a) complexed		
	Ligand	Δn_{exp}	Δn_{calc}	$\Delta n_{calc/HL}$	His57	ligCOO	ligAMINO	His57	ligCOO	ligAMINO
Trypsin	**1b**	0.90	0.51	0.49	0.43(7.68)	0.00(3.21)	n.a.	0.92(8.84)	0.00(2.69)	n.a.
	1bMe	0	−0.18	−0.14	0.20(7.20)	n.a.	n.a.	0.06(6.62)	n.a.	n.a.
	1c	0	−0.03	−0.01	0.20(7.21)	0.00(4.17)	n.a.	0.21(7.23)	0.00(3.69)	n.a.
	1cMe	0	−0.29	−0.21	0.30(7.44)	n.a.	n.a.	0.09(6.80)	n.a.	n.a.
	1d	−0.53	−0.14	−0.06	0.29(7.42)	n.a.	0.33(7.49)	0.00(5.50)	n.a.	0.56(7.90)
	1dAc	0	−0.21	−0.12	0.24(7.29)	n.a.	n.a.	0.12(6.92)	n.a.	n.a.
	2	0.93	0.57	0.54	0.46(7.73)	0.00(3.40)	n.a.	1.00(10.92)	0.00(3.30)	n.a.
	3	0	0.14	0.15	0.38(7.59)	0.00(3.84)	n.a.	0.53(7.85)	0.00(3.40)	n.a.
	4	0	0.34	0.36	0.38(7.58)	0.00(2.65)	0.32(7.48)	0.11(6.91)	0.00(0.77)	0.95(9.07)
	5	0	0.08	0.10	0.46(7.73)	0.00(2.51)	0.59(7.95)	0.18(7.13)	0.00(0.46)	0.97(9.28)
Thrombin	**4**	0	0.20	0.26	0.77(8.32)	0.00(2.65)	0.32(7.48)	0.37(7.56)	0.00(−0.68)	0.98(9.56)
	5	0	0.15	0.01	0.76(8.31)	0.00(2.51)	0.59(7.95)	0.36(7.55)	0.00(−0.85)	1.00(10.18)

The pK_a values of the titratable groups are given in parentheses. Δn_{calc} corresponds to the protonation effect considering all pK_a shifts observed in the 12 Å sphere; $\Delta n_{calc/HL}$ corresponds to the protonation effects considering the pK_a shifts of His57 and the ligand; ligCOO and ligAMINO refer to carboxylate and amino groups of the ligand. n.a., not applicable. See the text for further explanations.

are of limited significance if they are based on a pK_a shift of less than 1 unit (which is possible only at the critical buffer pH).

Rather small calculated pK_a shifts near the buffer pH for complexes with no measured protonation change might lead to the impression that a proton transfer is falsely estimated. However, such small shifts must be considered in the context of the limited significance mentioned above; e.g. a calculated protonation change of 0.33 mol based on a pK_a shift from 7.4 to 6.8 at pH 7.0 is not sufficient to postulate a true proton transfer.

Finally, the experimental error of the ITC measurement should be mentioned, which is estimated to be ±0.12 mol.[13] Such a protonation change corresponds to a pK_a shift of at least ±0.20 unit (if occurring at pK_a values around the buffer pH; larger, if occurring at pK_a values further from the buffer pH). This needs to be considered for the interpretation of the following results.

pK_a calculations for the complexes

For the binding of **1b** (Figure 3) to trypsin, ITC detected a proton uptake by the system of Δn_{exp} = +0.90. Visual inspection of the protein–ligand complex reveals either His57 or the carboxylate group of the ligand as likely candidates for picking up a proton. The distance between the N^{δ} of His57 and the Asp102 carboxy oxygen atom OD2 is 2.7 Å, and the N^{ε} is 3.7 Å remote from the ligand's carboxy oxygen atom (Figure 1(a)). Furthermore, the ligand's carboxylic function is within hydrogen bond distance (2.9 Å) from Ser195. In an earlier report, we assumed that the carboxy function of ligand **1b** would pick up the proton, simply because the corresponding methylester of **1b** did not show a similar protonation change.[13] The computed pK_a values of complexed and uncomplexed trypsin show a significant shift only for His57. For the apo structure, a pK_a value of 7.68 is calculated, whereas for the complex a pK_a of 8.84 is suggested. Taking all apparent pK_a shifts into account, a net uptake of $\Delta n_{calc/all}$ = +0.51 mol of protons is computed (96% of this proton uptake result from the His57 pK_a shift). In contrast, the ligand's carboxylic group shows no pK_a increase and, obviously, remains deprotonated.

In order to study whether the pK_a calculations would suggest unchanged protonation states for the binding of the ester **1bMe** to trypsin, a reasonable binding geometry of the ester had to be generated on the basis of the crystal structure of the acid **1b**. A crystal structure of the **1bMe** complex could not be obtained. Since both possibilities to attach the methyl group to the carboxylate oxygen atoms of **1b** appeared feasible, we performed calculations on the basis of both binding orientations. However, the two possibilities did not show major differences in the resulting pK_a values. The net change of protonation $\Delta n_{calc/all}$ is predicted to be −0.18.

Shifting the carboxylic group at the piperidine ring (as in **1b**) from the 2 to the 4 position gives **1c**. At first glance, this appears to be a minor variation. Both ligands **1b** and **1c** place their carboxylic functions in related regions of the binding pocket and, thus, similar protonation effects might be expected (Figure 1(a) and (b)). However, ITC revealed that no exchange of protons occurs upon complexation of **1c** (we call this a zero effect). On the basis of the calculations, the His57 pK_a is 7.21 in the uncomplexed form and remains at 7.23 in the complex. No further titratable group shows any significant pK_a shift that would give rise to a protonation change ($\Delta n_{calc/all}$ is −0.03).

Esterification of the ligand's 4-carboxylic function (**1cMe**) showed no protonation effect in the ITC experiment: Complexation of **1cMe** with trypsin does not result in proton uptake or proton release. The tendency for a zero effect is predicted correctly by our computations, although a net value $\Delta n_{calc/all}$ of −0.29 is suggested.

Upon binding of **1d** to trypsin, the complex shows an overall release of Δn_{exp} = −0.53 mol of protons in the ITC experiment. Visual inspection of the

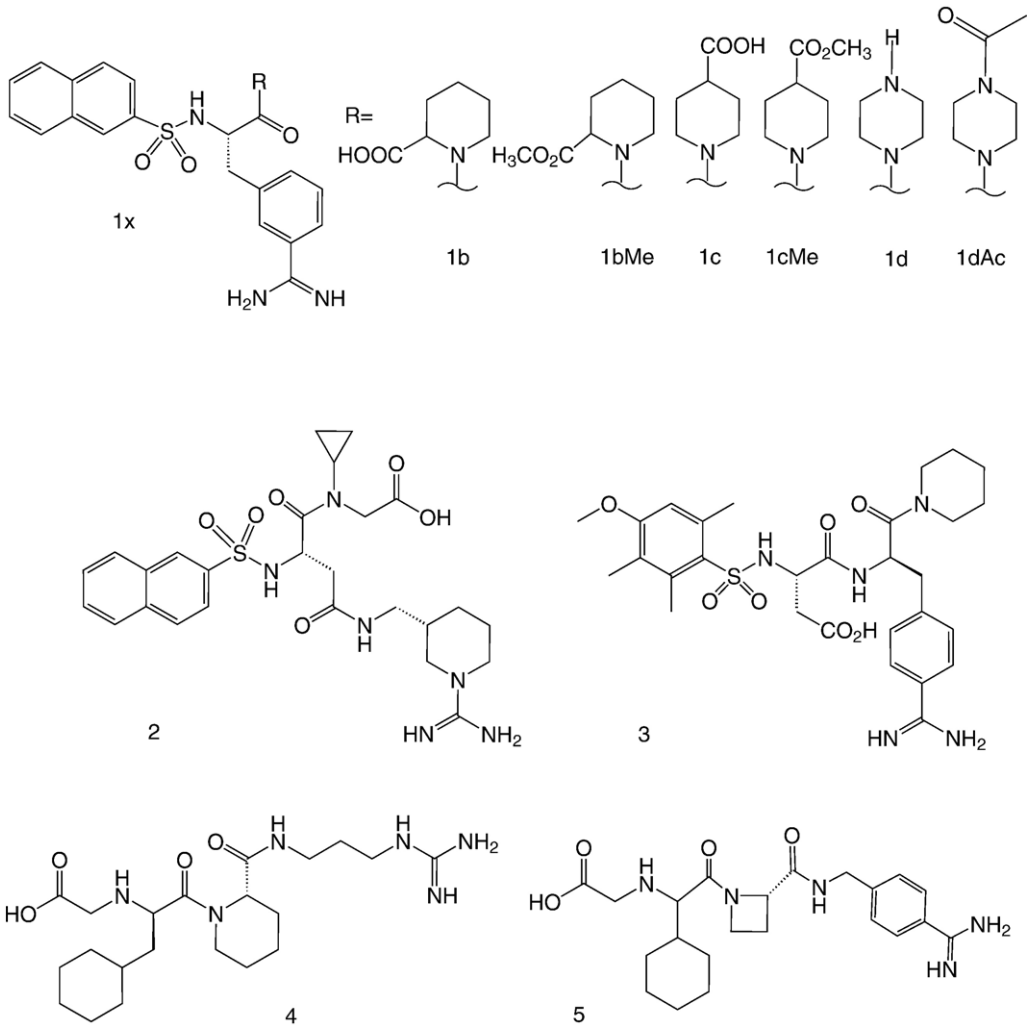

Figure 3. Chemical formulae of the ligands studied by Dullweber *et al.*[13] and used in this study, retaining the original numbering without changes.

complex geometry reveals a similar arrangement of the **1b.Try**, **1bMe.Try**, **1c.Try** and **1cMe.Try** complexes: The piperazine ring of **1d** occupies the same region of trypsin as in **1b**, **1bMe**, **1c**, and **1cMe** (Figure 1). However, a basic instead of an acidic function is now accommodated in the S_2 pocket. Although the conserved binding mode might imply a very similar situation in the **1c.Try** and **1d.Try** complexes, the altered functional group has dramatic influence on the pK_a calculation: It shows irregular (non-Henderson–Hasselbalch/non-sigmoidal) titration curves for His57 and the ligand. Such curves have been observed experimentally,[14] as well as theoretically (mostly in PB calculations).[15–18] Two requirements are necessary to produce irregular curves.[19] (1) The electrostatic interaction between the two groups has to be strong, and (2) they must titrate in the same pH range. In the case of the **1d.Try** complex, the two requirements are fulfilled: The strong interaction is caused by close proximity of the titrating groups; their distance amounts to 4.1 Å measured between the titrating nitrogen atoms. Furthermore, the two nitrogen functionalities titrate in the same pH range, since they are both basic and experience similar intrinsic pK_a values (the intrinsic pK_a value of a titratable group corresponds to its pK_a value when all other titratable groups are in the neutral state): the intrinsic pK_a values of His57 and the ligand are 5.38 and 7.07, respectively. In analogy to Trylska's treatment of the coupled system of the two aspartate residues in the catalytic dyad of HIV protease,[20] the two titration curves (His57/ligand) are superimposed and the His57/ligand piperazine system is considered as one coupled base system. Evaluating this titration curve, the following pK_a values can be extracted: His57 experiences a downward pK_a shift from 7.42 to 5.50, whereas the pK_a value of the ligand's secondary amino group changes from 7.49 (in aqueous solution) to 7.90 (in the complex). The resulting net protonation change (upon ligand binding) is estimated to be $\Delta n_{\mathrm{calc/all}} = -0.14$.

The experimentally observed deprotonation occurring upon **1d** complexation by trypsin can be prevented by acetylating the secondary nitrogen atom of the piperazine ring. This is evidenced also

by ITC measurement of the corresponding **1dAc.Try** complex. Our pK_a calculations yield a $\Delta n_{calc/all}$ of −0.21.

Ligand **2** differs structurally from the **1b–d** series (Figure 3). The pK_a calculation for the **2.Try** complex yields a protonation change of $\Delta n_{calc/all} = +0.57$ (95% of this Δn results from an upward pK_a shift of His57). The ITC experiment showed a net proton uptake of $\Delta n_{exp} = +0.93$.

Inhibitor **3** also bears a carboxylic function, but it differs structurally from the **1b–d** series (Figure 3). However, as the crystal structure of **3.Try** shows, its carboxylate group is hydrogen-bonded to Gln174, which is more than 7 Å apart from Ser195 and more than 8 Å from His57. Furthermore, the acidic group is strongly exposed to the adjacent solvent environment. ITC reveals no protonation effect upon binding of **3** to trypsin, the calculations give $\Delta n_{calc/all} = +0.14$.

Also the remaining two ligands for which pK_a calculations were performed differ structurally from the previous ligands. For both ligands, crystal structures with trypsin and thrombin have been determined. Interestingly, the ligands do not bear any aromatic side-chain to address the S_3 pocket, and they bind with their guanidino (**4**) or benzamidino group (**5**) into the S_1 pocket (Figure 1(d)). The ligands have two titratable groups in common: a terminal carboxylic group and a secondary amine, separated by two bonds. In the crystal structure, these groups are more than 9 Å remote from His57 and, thus, only marginal influence on its pK_a value is expected. Indeed, rather small pK_a shifts for His57 are observed in the pK_a calculations (see Table 2). The ITC experiment showed no proton exchange when **4** or **5** binds to trypsin or thrombin. The following overall changes in protonation were obtained (the values for thrombin are given in parentheses): $\Delta n_{calc/all}$ is +0.34 for **4.Try** (+0.20 for **4.Thr**), while for **5.Try** $\Delta n_{calc/all}$ is 0.08 (+0.15 for **5.Thr**).

Besides His57 and the ligand, there are further titratable groups that experience changes in pK_a values upon complex formation in all complexes considered. For example, Asp189 in the S_1 pocket of trypsin shows a downward pK_a shift (mean pK_a value for apo trypsin, 4.3; mean pK_a value for complexed trypsin, 2.6). Tyr228, also contributing to the S_1 pocket, is involved in another downward pK_a shift: it changes from 12.5 (mean pK_a value in apo trypsin) to 11.5 (mean pK_a value in complexed trypsin). However, these shifts do not provoke any change in overall protonation. Therefore, they can be neglected for the interpretation of the observed net release or uptake of protons.

Discussion

Our pK_a calculations provide evidence that the protonated His57 and deprotonated ligand **1b** form a salt-bridge-like electrostatic interaction. At first glance, this suggests that the placement of a charged carboxylate group next to the catalytic triad would generally induce a protonation change of His57. However, as **1c.Try** shows, this is not the case.

Closer inspection of the **1c.Try** complex shows that the ligand's carboxy oxygen atom (O55) is 5.1 Å apart from N^{ε} of His57, which corresponds to an expansion of 1.4 Å compared to **1b.Try**. In addition, the distance between the Ser195 oxygen atom O^{γ} and the ligand's oxygen atom (O55) is increased from 2.9 Å to 7.5 Å. Most likely, these longer distances alter the electrostatic properties of the complex (**1c.Try**) significantly. Accordingly, the pK_a calculation suggests no change in protonation for His57, which is in agreement with experiment. Purely on the basis of structural terms (without any pK_a calculation), it would be difficult to predict such a zero effect for the latter complex.

For the **1d.Try** complex, the results deviate from the experimental finding. Originally, it was assumed that a downward pK_a shift would occur only at the ligand's amino group upon trypsin binding.[13] The pK_a calculation suggests that His57 undergoes a significant downward pK_a shift, along with a rather marginal upward pK_a shift for the ligand. Although the ligand's pK_a value changes by less than 0.5 log unit, it almost outweighs the effect on His57. This example shows that a small pK_a shift in the physiological pH range may result in a very complex picture of mutually compensating effects that prevent a straightforward interpretation of experimentally observed protonation changes. The calculations suggest the correct trend for the net change in protonation of **1d.Try**, but the quantitative agreement is unsatisfactory. Nevertheless, assessing the relevance of such calculations, one has to bear in mind the experimental error range and the reduced accuracy in pK_a estimations that result from coupled titration curves. These are much more difficult to handle and allow only crude pK_a predictions.

The positions of the carboxylate group of ligand **2** in trypsin and thrombin are similar. Both complexes show almost equal protonation effects in the ITC experiment (**2.Thr** $\Delta n = +0.88$; **2.Try** $\Delta n = +0.93$). For both ligands, the esterified form has been complexed with trypsin and thrombin, and measured by ITC, and a zero effect was observed. Accordingly, it was originally assumed that the pK_a shift in **2.Thr** and **2.Try** occurs at the carboxylic group of the ligand.[13] In a previous computational study,[12] we could correctly quantify the effect for **2.Thr** using the available crystal structure. For **2.Try**, only a docked structure served as starting geometry for the pK_a calculation, since no corresponding crystal structure is available. Possibly, this explains the somewhat less satisfactory agreement of experiment and computer prediction for **2.Try** ($\Delta\Delta n = 0.36$) compared to **2.Thr** ($\Delta\Delta n = 0.28$).

The computer simulations suggest that the two titratable groups of **4** and **5** tend to adopt a zwitterionic state when complexed to the protein. The amino group and the carboxylate form a strongly coupled electrostatic system over short distance, which induces a downward pK_a shift for

the carboxylate group and an upward shift for the amino group. While the downward pK_a shift does not alter the protonation state of the carboxylic group, the upward pK_a shift of the amino function suggests a protonation change from the neutral to the positively charged state. Rather small pK_a shifts of the amino function are sufficient to provoke this net change; e.g. for **4** (p$K_{a(aq)}$ = 7.48), a pK_a shift of +0.5 unit induces a protonation change of +28% (−0.5 unit: −19%) referring to a buffered solution of pH 7.8. The net protonation inventory resulting from the pK_a shift of the amino groups in **4** and **5** might be partially compensated by a downward pK_a shift of His57, but still a value deviating from zero results as Δn_{calc}. Interestingly, if one of the two ligand titratable groups is considered as permanently charged and, thus, not involved in protonation changes, the measured zero effect is predicted much better by the simulations. This may indicate that the site–site interaction between the ligand titratable groups is overestimated in the pK_a calculation and prevents quantitative agreement. Nevertheless, for three of the four complexes, the zero effect is indicated correctly if we consider all Δn_{calc} values smaller than 0.2 as negligible (see sections Estimating the accuracy of calculated protonation changes and Conclusions).

Influence of conformational variations

To assess the influence of slightly different conformations on the computed pK_a values, we used the crystallographically determined apo structure of trypsin (1TGN) along with nine artificially generated uncomplexed structures (ligand-deleted trypsin, obtained by removing the bound ligand) to determine the pK_a values of uncomplexed trypsin. The calculated pK_a of His57 in 1TGN is 7.88. Analyzing all calculations on the nine ligand-deleted trypsin structures (Table 2) unveils a slight variation of the pK_a for the crucial residues His57 (7.49 ± 0.19) and Asp102 (1.99 ± 0.36). This suggests that the uncertainties in our calculations will be at least in the range of the estimated experimental error of ITC measurements and provides an additional approximation of the significance of the computed pK_a shifts. The strategy of using the ligand-deleted trypsin structures appeared justified, as Guvench *et al*. showed by MD simulations that a high level of structural conservation of the trypsin-binding pocket is given.[21] The overall RMSD based on a C^{α}-fit between 1TGN and our trypsin complexes ranges between 0.28 Å and 0.42 Å, which illustrates the low degree of flexibility of the trypsin structures in response to ligand binding.

pK_a values of histidine residues

In the present study, pK_a shifts of His57 are crucial for the protonation inventory. Therefore, it is instructive to examine the spread of experimentally observed pKa shifts of histidine residues in proteins, also as an example to estimate the impact of the protein environment on pK_a values. While in our calculations for the different complexes, the pK_a of His57 varies over a range of approximately 5 units, the experimental pK_a values of histidine residues in various protein structures are reported to be distributed over more than 7 units.[22] His149 in *Bacillus circulans* shows a value below 2.3.[23] In contrast, the pK_a values of His66 and His72 in a protein tyrosine phosphatase are 8.3 and 9.2, respectively.[24] The pK_a value of His31 in bacteriophage T4 lysozyme is 9.1, caused by the formation of a salt-bridge to Asp70 (pK_a = 0.5).[25] More than 300 pK_a values for histidine residues are available in a recently published data base.[26] The mean value is 6.33 (standard deviation of 1.35); and 35 deposited histidine residues exhibit a pK_a value greater than 7.5. The experimental pK_a of His57 in rat trypsin is 6.7 (NMR measurement for a D189S mutant),[27] whereas the pK_a value of the imidazole moiety of the free amino acid in water has been determined to be 6.3.[28]

Conclusions

In this study, systematic investigations of protonation changes upon ligand binding to trypsin and thrombin have been performed. For nine complexes that show a zero effect in ITC, this protonation inventory was predicted correctly, and rather small $\Delta n_{calc/all}$ values have been computed: −0.18 (**1bMe.Try**), −0.03 (**1c.Try**), −0.29 (**1cMe.Try**), −0.21 (**1dAc.Try**), +0.14 (**3.Try**), +0.34 (**4.Try**), +0.08 (**5.Try**), +0.20 (**4.Thr**), +0.15 (**5.Thr**). On average, a net change of ±0.18 mol of protons (with a standard deviation of 0.09) is suggested by the calculations for these complexes. These minor shifts have to be seen in light of the experimental accuracy for protonation changes, which is estimated to be Δn = ±0.12.[13] The computed predictions deviate more for complexes where most likely a coupled system of titratable groups is given. The behaviour of such groups deviates from a simple Henderson–Hasselbalch description, and thus complicates a straightforward handling. Furthermore, caution is warranted if minor pK_a shifts near the buffer pH are predicted, as they still translate into rather large changes of the protonation inventory (see the section Estimating the accuracy of calculated protonation changes).

In cases where ITC suggests a net change in the protonation inventory upon complex formation, our calculations indicate the correct trends. For one complex (**1d.Try**), only the qualitatively correct tendency is suggested; however, this complex seems to exhibit strongly coupled titratable groups.

The present calculations allow structural interpretation of protonation states along with ITC measurements. ITC experiments give access to the Gibbs free energy ΔG along with the heat absorbed or produced during complex formation. This heat contains the binding enthalpy, but changes in protonation states may be superimposed. This is due to the fact that any release or pick-up of protons

involves a heat of ionization of the functional groups that participate in the protonation exchange process. In consequence, it concerns the functional groups of the buffer molecules as well as the groups of the protein residues or the bound ligand that change their protonation state. In this context, it matters dramatically whether the groups involved in the protonation reaction exhibit oxygen or nitrogen functionalities.

In earlier work, we wrongly assumed that the carboxylate function of **1b** and **2** would undergo protonation upon trypsin and thrombin binding.[13] This interpretation was born out of the simple conclusion that the corresponding esters of **1b** and **2** did not show any protonation change. Accordingly, we corrected for a heat of ionization contribution of 2.26 kJ/mol or 0.84 kJ/mol, respectively, for the acidic oxygen functions.[13] After subtraction of this contribution from the measured heat, a net ΔH_{bind} of −45.6 kJ/mol for **1b.Try** was assumed. The measured ΔG_{bind} was then factorized into enthalpy and entropy. The calculations presented here suggest that the carboxylic function of the ligand remains deprotonated, whereas His57 picks up a proton. This now involves ionization of a nitrogen functionality that produces a much higher heat contribution (ΔH_{ion} of His, 30 kJ/mol†). Accordingly, ionization of His57 upon complex formation of **1b** with trypsin requires a correction of 27 kJ/mol (0.90 × 30 kJ/mol). After subtraction of this value from the measured heat, a quite different value for $\Delta H_{bind} = -20.6$ kJ/mol remains. In consequence, this suggests a strongly deviating factorization in enthalpic and entropic contributions. Similar corrections have to be performed for the complexes of **2** (see Table 3).

Accordingly, the interpretation of the driving forces responsible for binding of **1b** and **2** to the considered serine proteases has to be modified (Table 3). Binding is now characterized by a significantly smaller enthalpic and a more favourable entropic contribution. The originally suggested pronounced enthalpic binding of **2** to thrombin is now reduced and, thus, is more comparable with data of the other ligands in the series. Furthermore, the formerly contrasting partitioning for the complexes of acid **1b** and ester **1bMe** with trypsin is now suggested to be of equal size for both inhibitors. Even though the stronger binding of napsagatran (**2**) to thrombin results (still) from a stronger exothermic binding compared to the ester, the entropic advantage of the ester compared to the acid is much less pronounced according to the new partitioning.

For **1d.Try**, a more complex protonation transfer is suggested by the computer simulations. However, with respect to the partitioning of enthalpy and entropy, the assignment of heat of ionization corrections remains virtually unchanged, because the amount to be corrected for the histidine residue is very similar to that of the ligand's piperazine moiety. Accordingly, no significant change with respect to

† http://www.lsbu.ac.uk/biology/enztech/ph.html

Table 3. Corrected factorization of enthalpic and entropic binding contributions from ITC experiments

Complex	Δn_{obs} Molar ratio of protons transferred	$K(10^6\ M^{-1})$ for trypsin, K_i (nM) for thrombin	ΔG free energy derived from binding constant	ΔH_x enthalpy corrected only for buffer contribution	ΔH^{COO} molar ionization enthalpy of carboxylate	ΔH^{His} molar ionization enthalpy of imidazole	ΔH^{old}_{bind} former corrected enthalpy	ΔH^{new}_{bind} current corrected enthalpy	$-T\Delta S^{old}$ former corrected entropy	$-T\Delta S^{new}$ current corrected entropy
1b.Try	0.90±0.06	2.44±0.19	−36.4±0.2	−47.6±1.9	2.0[a]	30.0[a]	−45.6±1.9	−20.6±1.9	9.2±1.9	−15.8±1.9
1bMe.Try	0	3.11±1.10	−37.0±1.1	−16.9±0.8	Not necessary	Not necessary	−16.9±0.8	−16.9±0.8	−20.6±1.3	−20.6±1.3
2.Try	0.93±0.12	0.40±0.07	−32.0±0.5	−32.3±3.6	0.8[a]	30.0[a]	−31.5±3.6	−4.4±3.6	−0.5±3.6	−27.6±3.6
2.Thr	0.88±0.08	0.70[a]	−52.2	−75.1±1.8	0.8[a]	30.0[a]	−74.4	−48.7	22.2	−3.5
2Et.Thr	0	86[a]	−40.3	−24.4±0.9	Not necessary	Not necessary	−24.4	−24.4	−15.9	−15.9

All energy values are given in kJ/mol. Experimental errors are indicated for the directly measured values. Error propagation has been applied for the derived values.
[a] No experimental error given in literature (accordingly, no error propagation could be applied).

our previous interpretation and factorization is required for **1d.Try**.

In summary, protonation changes can lead to very different factorization into enthalpy and entropy, depending on the ionizable groups involved. The pK_a calculations presented here support the thermodynamic characterization and interpretation of the protein–ligand binding process, if protonation effects are superimposed. The experimental efforts to fully characterize protonation changes are quite elaborate and consume large amounts of expensive protein material. In contrast, the calculations described here are cheap and fast to perform. As such, they can provide valuable assistance in properly planning and analyzing ITC experiments.

Materials and Methods

The investigated ligands are shown in Figure 3. Most of them contain three titratable groups: a guanidino group, a sulfonamide or secondary amino group, and a carboxylic acid group. Experimental pK_a values have been determined for all ligands.[13] In the present pK_a calculations, only the carboxylic acid and secondary amino group were considered as titratable, since their pK_a values fall into the physiological pH range. A 3D representation of the complexes is shown in Figure 1. The following complex crystal structures (referred to by their PDB code) were used as starting geometries: 1K1I (**1b.Try**), 1K1J (**1cMe.Try**), 1K1L (**1d.Try**), 1K1M (**1dAc.Try**), 1K1N (**3.Try**), 1K1O (**4.Try**), 1K1P (**5.Try**); for thrombin, the complexes 1K21 (**4.Thr**) and 1K22 (**5.Thr**) were used. A geometry of the trypsin complex with **1c** was generated by replacing the methylester of **1cMe.Try** with a carboxylic acid group. Subsequently, the modelled complex was energy-minimized with the program MOLOC using the MAB forcefield.[29] The geometry of the complex **1bMe.Try** was generated similarly but in reverse fashion by attaching a methyl group to either oxygen atom of the carboxylic function of **1b**; thus, both possibilities for ester formation have been explored. A 3D structure of the **2.Try** complex was obtained by docking using GOLD.[30] 1K1P was chosen as template structure, and 20 docking poses were generated and evaluated. The rank 3 solution was selected for the pK_a calculation on the basis of its convincing agreement with the thrombin/napsagatran crystal structure.[13]

The pK_a calculations were done as described.[12] To study the protonation effects upon ligand binding, two pK_a calculations are required: one without ligand and a second with the bound ligand. MEAD was employed as PB solver (using a dielectric constant of 20 for the protein),[31] REDUCE was used to generate hydrogen atoms for the residues of the protein.[32] SYBYL was applied to add hydrogen atoms to the ligands‡. Only the titratable residues within a 12 Å sphere around the active site were selected for evaluation of the site–site interactions in order to reduce the computational effort (in total, there are 39 titratable residues in trypsin and 81 in thrombin). This results in the following titratable residues for trypsin: Tyr39, His40, His57, Tyr59, Lys60, Tyr94, Asp102, Lys145, Tyr151, Tyr172, Asp189, Asp194, Lys224, Tyr228. In the case of thrombin, the following titratable residues lie within the sphere: Glu18, Glu39, Asp56E, His57, Tyr60A, Lys60F, Tyr94, Glu97A, Asp100, Asp102, Lys145, Glu146, Lys169, Tyr184A, Asp189, Glu192, Asp194, Glu217, Asp221, Asp222, Lys224, Tyr225, Tyr228.

‡SYBYL Molecular Modeling Software, 7.0 ed.; Tripos Inc., St. Louis, MO, USA.

Acknowledgement

The authors gratefully acknowledge financial support by the bilateral CERC3 program of CNRS and DFG (KL 1204/3).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2007.01.022

References

1. Bashford, D. & Karplus, M. (1990). pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry*, **29**, 10219–10225.
2. Antosiewicz, J., McCammon, J. A. & Gilson, M. K. (1994). Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* **238**, 415–436.
3. Nielsen, J. E. & Vriend, G. (2001). Optimizing the hydrogen-bond network in Poisson-Boltzmann equation-based pka calculations. *Proteins: Struct. Funct. Genet.* **43**, 403–412.
4. Nielsen, J. E. & McCammon, J. A. (2003). Calculating pKa values in enzyme active sites. *Protein Sci.* **12**, 1894–1901.
5. Warshel, A. (1981). Calculations of enzymic reactions: calculations of pKa, proton transfer reactions, and general acid catalysis reactions in enzymes. *Biochemistry*, **20**, 3167–3177.
6. Warshel, A. & Levitt, M. (1976). Theoretical studies of enzymatic reactions: dielectric electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **103**, 227–249.
7. Warshel, A. & Russell, S. T. (1984). Calculations of electrostatic interactions in biological systems and in solutions. *Quart. Rev. Biophys.* **17**, 283–422.
8. Sham, Y. Y., Chu, Z. T. & Warshel, A. (1997). Consistent calculations of pKa's of ionizable residues in proteins: semi-microscopic and macroscopic approaches. *J. Phys. Chem. B*, **101**, 4458–4472.
9. Li, H., Robertson, A. D. & Jensen, J. H. (2005). Very fast empirical prediction and rationalization of protein pKa values. *Proteins: Struct. Funct. Genet.* **61**, 704–721.
10. Wisz, M. S. & Hellinga, H. W. (2003). An empirical model for electrostatic interactions in proteins incorporating multiple geometry-dependent dielectric constants. *Proteins: Struct. Funct. Genet.* **51**, 360–377.
11. Mehler, E. L. & Guarnieri, F. (1997). A self-consistent, microenvironment modulated screened coulomb potential approximation to calculate pH-dependent electrostatic effects in proteins. *Biophys. J.* **77**, 3–22.
12. Czodrowski, P., Dramburg, I., Sotriffer, C. A. & Klebe, G. (2006). Development, validation and application of adapted PEOE charges to estimate pKa

values of functional groups in protein-ligand complexes. *Proteins: Struct. Funct. Genet.* **65**, 424–437.
13. Dullweber, F., Stubbs, M. T., Musil, D., Stürzebecher, J. & Klebe, G. (2001). Factorising ligand affinity: a combined thermodynamic and crystallographic study of trypsin and thrombin inhibition. *J. Mol. Biol.* **313**, 593–614.
14. Kesvatera, T., Jonsson, B., Thulin, E. & Linse, S. (1999). Ionization behavior of acidic residues in calbindin. *Proteins: Struct. Funct. Genet.* **37**, 106–115.
15. Yang, A. S., Gunner, M. R., Sampogna, R., Sharp, K. & Honig, B. (1993). On the calculation of pKas in proteins. *Proteins: Struct. Funct. Genet.* **15**, 252–265.
16. Onufriev, A., Case, D. A. & Ullmann, G. M. (2001). A novel view of pH titration in biomolecules. *Biochemistry*, **40**, 3413–3419.
17. Bashford, D. & Gerwert, K. (1992). Electrostatic calculations of the pKa values of ionizable groups in bacteriorhodopsin. *J. Mol. Biol.* **224**, 473–486.
18. Carlson, H. A., Briggs, J. M. & McCammon, J. A. (1999). Calculation of the pKa values for the ligands and side chains of *Escherichia coli* D-alanine:D-alanine ligase. *J. Med. Chem.* **42**, 109–117.
19. Klingen, A. R., Bombarda, E. & Ullmann, M. (2006). Theoretical investigation of the behavior of titratable groups in proteins. *Photochem. Photobiol. Sci.* **6**, 588–596.
20. Trylska, J., Antosiewicz, J., Geller, M., Hodge, C. N., Klabe, R. M., Head, M. S. & Gilson, M. K. (1999). Thermodynamic linkage between the binding of protons and inhibitors to HIV-1 protease. *Protein Sci.* **8**, 180–195.
21. Guvench, O., Price, D. J. & Brooks, C. L. (2005). Receptor rigidity and ligand mobility in trypsin-ligand complexes. *Proteins: Struct. Funct. Genet.* **58**, 407–417.
22. Edgecomb, S. & Murphy, K. (2002). Variability in the pKa of histidine side-chains correlates with burial within proteins. *Proteins: Struct. Funct. Genet.* **49**, 1–6.
23. Plesniak, L. A., Connelly, G. P., Wakarchuk, W. W. & McIntosh, L. P. (1996). Characterization of a buried neutral histidine residue in bacillus circulans xylanase: NMR assignments, pH titration, and hydrogen exchange. *Protein Sci.* **5**, 2319–2328.
24. Tishmack, P. A., Bashford, D., Harms, E. & Van Etten, R. L. (1997). Use of 1H NMR spectroscopy and computer simulations to analyze histidine pKa changes in a protein tyrosine phosphatase: experimental and theoretical determination of electrostatic properties in a small protein. *Biochemistry*, **36**, 11984–11994.
25. Anderson, D. E., Becktel, W. J. & Dahlquist, F. W. (1990). pH-induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry*, **29**, 2403–2408.
26. Toseland, C. P., McSparron, H., Davies, M. N. & Flower, D. R. (2006). PPD v1.0–an integrated, web-accessible database of experimentally determined protein pKa values. *Nucl. Acids Res.* **34**, 199–203.
27. Kaslik, G., Westler, W. M., Graf, L. & Markley, J. L. (1999). Properties of the His57-Asp102 dyad of rat trypsin D189S in the zymogen, activated enzyme, and alpha1-proteinase inhibitor complexed forms. *Arch. Biochem. Biophys*, **362**, 254–264.
28. Tanokura, M. (1983). 1H-NMR study on the tautomerism of the imidazole ring of histidine residues: I. Microscopic pK values and molar ratios of tautomers in histidine-containing peptides. *Biochim. Biophys. Acta*, **742**, 576–585.
29. Gerber, P. & Müller, K. (1996). MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput. Aid. Mol. Des.* **9**, 251–268.
30. Jones, G., Willett, P. & Glen, R. C. (1995). Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **245**, 43–52.
31. Bashford, D. (1997). An object-oriented programming suite for electrostatic effects in biological molecules. In *Scientific Computing in Object-oriented Parallel Environments* (Ishikawa, Y., Oldehoeft, R. R., Reynders, J. V. W. & Tholburn, M., eds), vol. 1343, pp. 233–240 Springer, Berlin.
32. Word, J. M., Lovell, S. C., Richardson, J. S. & Richardson, D. C. (1999). Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* **285**, 1735–1747.

Edited by J. Ladbury

(Received 25 August 2006; received in revised form 19 December 2006; accepted 5 January 2007)
Available online 12 January 2007