Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 25 (2006) 176–185
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Predicting proteinase specificities from free energy calculations

Seble Merid Mekonnen, Magne Olufsen, Arne O. Smalås, Bjørn O. Brandsdal *

The Norwegian Structural Biology Centre, Faculty of Science, University of Tromsø, N9037 Tromsø, Norway

Received 28 August 2005; received in revised form 11 November 2005; accepted 14 November 2005
Available online 10 January 2006

Abstract

The role of the primary binding residue (P1) in complexes between three different subtilases (subtilisin Carlsberg, thermitase and proteinase K) and their canonical protein inhibitor eglin c have been studied by free energy calculations. Based on the crystal structures of eglin c in complex with subtilisin Carlsberg and thermitase, and a homology model of the eglin c–proteinase K complex, a total of 57 mutants have been constructed and docked into their host proteins. The binding free energy was then calculated using molecular dynamics (MD) simulations combined with the linear interaction energy (LIE) method for all complexes differing only in the nature of the amino acid at the P1 position. LIE calculations for 19 different complexes for each subtilase were thus carried out excluding proline. The effects of substitutions at the P1 position on the binding free energies are found to be very large, and positively charged residues (Arg, Lys and His) are particularly deleterious for all three enzymes. The charged variants of the acidic side chains are found to bind more favorably as compared to their protonated states in all three subtilases. Furthermore, hydrophobic amino acids are accommodated most favorably at the S1-site in all three enzymes. Comparison of the three series of binding free energies shows only minor differences in the 19 computed relative binding free energies among these subtilases. This is further reflected in the correlation coefficient between the 23 relative binding free energies obtained, including the possible protonation states of ionizable side chains, but excluding the P1 Pro, for subtilisin Carlsberg versus thermitase (0.95), subtilisin versus proteinase K (0.94) and thermitase versus proteinase K (0.96).

Keywords: Molecular dynamics simulations; Serine proteinases; Free energy calculations; Protein specificity

1. Introduction

The interactions between the canonical standard mechanism [1,2] serine proteinase inhibitors and their cognate enzymes [3,4] represent excellent models for the study of protein–protein recognition. Both substrate and inhibitor recognition by most serine proteinases involve accommodation of the primary binding residue (P1, notation of Schechter and Berger [5], see Fig. 1) to the specificity pocket (S1-site) of the enzyme. During this process the side chain of the P1 residue becomes buried in the S1 cavity of the enzyme, and the nature of the amino acid at this position greatly affects both the strength and the specificity of the non-covalent association.

Complex formation of serine proteinases with their natural inhibitors usually involves more than 20 residues from the proteinase and in the order of 10–17 from the protein inhibitor. The loop segment between P3 and P3′ of the inhibitor constitutes the principal contact area. This binding loop is very rigid and is not altered upon complex formation, and the primary recognition site is the solvent exposed P1 residue. The side chain of this residue is of particular importance for the proteinase inhibitor association energy and in most natural subtilase inhibitors, the P1 residue is Leu whereas trypsin inhibitors have Arg/Lys, chymotrypsin inhibitors have Leu/Met and in porcine pancreatic elastase inhibitors, it is Leu/Ala. In many inhibitor families the amino acid at the P1 position is recognized by a high degree of variability. As an example, 11 different amino acid residues were found at the P1 position in 125 reported sequences of ovomucoid third domains [6]. This variability resulted in different spectra of proteinases inhibited by the respective inhibitors.

Eglin c is a 70 amino acid residue serine proteinase inhibitor of potato I family, isolated from the leech *Hirudo medicinalis* [7], where residues 40–49 constitute the binding loop involved in enzyme substrate interactions [8]. It is a well-known model of a canonical inhibitor where both the X-ray [9] and NMR [10] structures have been determined for the free inhibitor, along with X-ray structures of complexes with different serine proteinases: subtilisin Carlsberg [11], chymotrypsin [12]. Eglin c has no

* Corresponding author. Tel.: +47 77 64 40 57; fax: +47 77 64 47 65.
E-mail address: Bjorn-Olav.Brandsdal@chem.uit.no (B.O. Brandsdal).

1093-3263/$ – see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2005.11.005

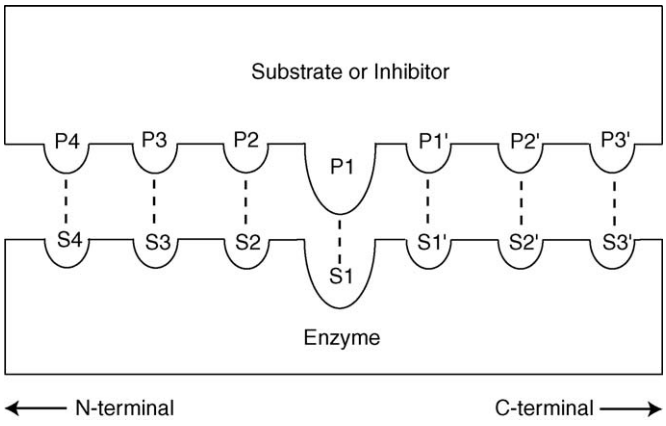

Fig. 1. Schematic diagram showing the interaction model introduced by Schechter and Berger [5] to describe the interaction between serine proteinases and their substrates or inhibitors.

disulfide bridges and is thus stabilized solely by non-covalent interactions.

Based on sequence homology, the subtilases are divided into six groups, i.e. the subtilisin, thermitase, proteinase K, lantibiotic peptidase, kexin and pyrolysin families [13]. Due to the great applicability as protein degrading agents in a wide range of biotechnological, food processing and other industrial processes, subtilases from psychrophilic [14,15], mesophilic [16] and thermophilic [17,18] origin have been extensively studied to gain insight into their thermostability [19,20], adaptability [21,22], catalytic activity [23,24] and substrate specificity [25,26].

A more thorough insight into the molecular origin of the variation in, e.g., specificity among subtilases will, however, make them more interesting as targets for protein engineering studies and thereby also increase their application potential. Efficient methods for screening of specificity profiles, coupled with structural studies of complexes, are needed in order to establish definite structure–specificity relation rules for a wide range of different subtilases. In this study we have shown that theoretical calculations with the linear interaction energy (LIE) method [27] can be used to investigate the interaction energies in complexes between subtilases and proteinase inhibitors. The LIE method was used to investigate specificity profiles for the subtilase–eglin c complexes of subtilisin Carlsberg, thermitase and proteinase K. The main objective of this work was to establish methods able to reproduce experimental data on protein–protein interactions with subtilases. The choice of targets for the calculations was therefore based on the availability of three-dimensional structures of subtilases and subtilase-inhibitor complexes, and corresponding association constant measurements. If the method is able to reproduce the effect of point mutations, then it may also be applicable for specificity screening of more general systems with native and in silico engineered subtilase–substrate complexes. Such data may in turn be used as a basis for in vitro mutational experiments aimed at design and redesign of both primary and secondary binding sub-site specificity of subtilases.

2. Methods

2.1. Protein–protein models

The two crystal structures of subtilisin Carlsberg (pdb entry 1CSE [11]) and thermitase (pdb entry 2TEC [28]) in complex with the protein inhibitor eglin c were used as templates in order to generate the series of 19 different complexes differing only in the amino acid at the P1 position. Proline was left out due to geometric considerations imposed by the side chain. Native eglin c has Leu at the primary binding position, and the remaining 18 complexes were generated using the crystallographic software O [29]. Near contacts and overlapping water molecules were manually removed. The third protein investigated was proteinase K, but in this case no experimental structure of the protein bound to eglin c was available to be used as a template. However, the proteinase K has a sequence identity of 60% compared to subtilisin Carlsberg, and one can therefore assume that the binding mode with eglin c is similar. The proteinase K (pdb entry 1IC6 [30]) was therefore superimposed onto the subtilisin Carlsberg complex to obtain a model of the protein–protein complex, which was subsequently used as a template to generate the remaining 18 complexes. The protein–protein interface was investigated and side chains were adjusted to remove near contacts. Steric clashes and overlapping water molecules (protein water contacts) were again manually removed by adjusting the amino acid side chains at the interface using the O software. These models were additionally improved through a MD simulation at 1 K, corresponding to a steepest decent minimization, prior to the heating phase as described below.

2.2. Molecular dynamics simulations

After the three series of protein–protein complexes had been built, MD simulations of each complex were carried out with the MD program package Q [31] using the Amber95 force field [32]. The Cα-atom of the P1 residue was defined as the simulation center, and atoms within 18 Å were allowed to move during the simulations. Atoms in the outermost 3 Å of this sphere were subjected to a harmonic positional restraint of 5 kcal/(mol Å^2), while those outside the 18 Å simulation sphere were heavily restrained to their initial positions. The nonbonded potentials were truncated at 10 Å, except for the P1 residue that were allowed to interact with all atoms in the simulation sphere. Interactions beyond the 10 Å cutoff were included using a multipole expansion [33]. Only bonded interactions across the boundary are included, which results in a significant reduction in the evaluation of nonbonded interactions. Water molecules closer than 12 Å from the simulation centre were included in the simulations. These are often highly conserved in serine proteinases and can play an important role in mediating binding specificity, and were therefore considered explicitly. Additional solvent molecules were added to fill the 18 Å simulation sphere, and all water

Table 1
Residues described with a charged topology in the MD simulations

Structure	Positive charges[a,b,c]	Negative charges[a,b,c]	Net charge
Subtilisin Carlsberg–eglin c	His 63 His 66 Arg 315 Arg 318 Arg 320 His 335	Asp 32 Asp 59 Asp 313 GCT 337	+2
Eglin c	Arg 315 Arg 318 Arg 320 His 335	Asp 313 GCT 337	+2
Thermitase–eglin c	His 71 His 74 KNT 280 Arg 320 Arg 323 Arg 325 His 340	Asp 38 Asp 105 Asp 318 GCT 342	+4
Eglin c	KNT 280 Arg 320 Arg 323 Arg 325 His 340	Asp 318 GCT 342	+4
Proteinase K–eglin c	His 69 His 72 Arg 320 Arg 323 Arg 325 His 340	Asp 39 Asp 97 Asp 318 GCT 342	+2
Eglin c	Arg 320 Arg 323 Arg 325 His 340	Asp 318 GCT 342	+2

[a] The numbering used is subtilisin BPN′ numbering.
[b] GCT denotes the glycine C-terminal of the inhibitor with charge −1.
[c] KNT denotes the lysine N-terminal of the inhibitor with charge +2.

molecules were described with the TIP3P model [34]. The molecular systems were heated stepwise from 1 to 300 K prior to an equilibration period to stabilize the P1-surrounding interaction energies. A timestep of 1.5 fs was used and energies were sampled every fifth integration step during the production phase. The production phase consisted of at least 225 ps, but in some cases additional simulation time was needed in order to obtain converged energies (up 450 ps). The criteria used to determine convergence was that the average interaction energy between the P1 residue and its surroundings should be similar in the first and second half of the simulation. If these two averages differed by more than 3.0 kcal/mol additional steps were included.

To avoid possible Born-terms from entering into the calculated binding free energy of ionic P1 variants, special care were taken with regard to the ionizable amino acids included in the 18 Å sphere. As has been discussed previously [35,36] keeping the total charge within the simulation sphere of the complex and of the inhibitor in solution the same ensures that the Born-terms cancel out. This can be achieved in two ways, either by adding counterions or by using neutral descriptions of amino acids close to the spherical boundary. In our simulations, we have used the latter method as addition of counterions may lead to convergence problems [36]. Residues that are described as being charged are given in Table 1. The contribution from residues that are described as neutral is added to the binding free energies by using a Columbic potential with a high dielectric constant, as discussed previously [35,36].

2.3. The linear interaction energy method

The linear interaction [27] method was used to calculate the free energy of binding between the protein–protein complexes based on the energetics from the MD simulations. This method has recently been reviewed [37,38] and only the basic principles will be presented here. The key idea is that the binding free energy can be estimated as a sum of a polar and a nonpolar contribution. Linear response theory is used to determine the polar part while the nonpolar part is estimated according to an empirically determined relationship. Only the physical relevant states are considered, that is the ligand bound to the proteins active site and when it is free in solution. Molecular dynamics (MD) or Monte Carlo (MC) simulations are used to generate ensembles of these two states, and average ligand–surrounding interaction energies are calculated. These energies are then used in the LIE equation:

$$\Delta G_{\text{bind}}^{\text{LIE}} = \alpha\,\Delta\langle V_{\text{l-s}}^{\text{vdw}}\rangle + \beta\,\Delta\langle V_{\text{l-s}}^{\text{elec}}\rangle + \gamma$$

where the Δ's denote changes in the ligand–surrounding energies between the bound and unbound state. $\langle V_{\text{l-s}}^{\text{elec}}\rangle$ and $\langle V_{\text{l-s}}^{\text{vdw}}\rangle$ are the average ligand–surrounding (l–s) interaction energies from the simulations. α and β are the scaling coefficients for the nonpolar and polar contribution, respectively, while γ is a coefficient needed in some systems to reproduce absolute binding free energies but not relative [39]. The polar coefficient assumes values depending on the chemical nature of the ligand ranging from 0.33 to 0.50 according to the number of hydroxyl groups in the ligand and whether the ligand is charged or not [40]. Here, charged amino acids use 0.50 and amino acids with hydroxyl groups use 0.37, while the remaining use 0.43. In contrast, the nonpolar coefficient is derived by optimization with respect to reproduction of experimental binding free energies. A recent study [41] has shown that a value of approximately 0.58 is necessary in protein–protein systems, while a value of 0.18 is normally used when the ligands are of more drug-like character [37,38]. The higher nonpolar coefficient used in protein–protein interactions primarily reflects the simplification of the linear response approximation used in order to estimate the electrostatic contribution in the LIE framework. The linear response theory requires four simulations to be carried out to evaluate the polar or electrostatic contribution to the binding free energy [27,42], corresponding to two physical states of the ligand (in water and in the binding site of the protein) and two unphysical states (the ligand in water and in the protein but now with its atomic charges turned off). The simulations for which the atomic charges are turned off are not carried out in the LIE context, but rather included through optimization of the nonpolar constant as discussed in Almlöf et al. [41]. In this study, optimization with respect to minimizing the error between the experimental and the calculated binding free energies relative to P1 Gly yields a nonpolar coefficient of 0.58 for subtilisin Carlsberg with a ligand dependent β.

3. Results and discussion

The specificity of serine proteinases is to a large extent determined by the shape, size and charge of the specificity pocket (S1-site), and the interaction between serine proteinases and their

standard mechanism canonical protein inhibitors has for a long time served as a classical example to understand protein specificity in general. Such standard mechanism protein inhibitors bind to the cognate proteinase in a substrate-like manner, but unlike natural substrates, form highly stable complexes with a lifetime that enables us to study the specific details involved in protein recognition. Several protein inhibitors have been subjected to site-directed mutagenesis and combined with kinetic, thermodynamic and structural studies to reveal the details on binding and complex formation with their cognate proteinases. However, such experimental procedures are very time consuming as well as expensive, and it is therefore desirable to have accurate theoretical approaches covering model building of the bound state with a subsequent determination of the energetic details underlying complex formation. This would, of course, open the possibility that the effect of point mutations on association energies can be rapidly determined, and provide a 'fingerprint' of the protein specificity.

Here the specificity profile for three different serine proteinases has been investigated with computer simulations. In a recent study by Almlöf et al. [41] the effect of point mutations at the P1 position on the association energy between chymotrypsin and elastase with OMTKY3 and trypsin with BPTI was investigated with the LIE method with success. Furthermore, the optimal coefficients in the LIE equation were found to be very similar in all three systems studied [41]. The idea of treating only the P1 residue as the 'ligand' in the LIE framework is very appealing as the interaction energy between this residue and its surrounding environment converges very quickly in the MD simulations. This would of course not be the case if the entire protein was treated as a ligand, as the interaction energy involved in protein–protein recognition that needs to converge is often on the order of several thousands kcal/mol. When applying the LIE method to study the effect of specific mutations on association energies it is very convenient to choose a reference state and relate the interaction energy of the remaining amino acids to this state. The reference state chosen can be somewhat arbitrary, but arguments have been presented that Gly is a particularly good choice [43]. Using P1 Gly, which does not have any side chain as a reference state, the relationship between the P1 specificity of the enzymes can be estimated according to the $\Delta\Delta G$ values. This $\Delta\Delta G(\mathrm{Gly} \rightarrow \mathrm{X}) = \Delta G_{\mathrm{X}} - \Delta G_{\mathrm{Gly}}$ can be taken as a quantitative measure of the interaction between the P1 X side chain and the specificity pocket of the cognate enzyme. If all the secondary binding contacts between the various P1 mutants are identical, the contribution to the $\Delta\Delta G$ is primarily from the P1–S1 interaction. However, if the secondary interactions change as a response to mutations at the P1 position, this will not be reflected in the computed binding free energies. This is observed when introducing large side chains into the shallow specificity pocket of elastase [41], but this can be easily addressed by investigating the corresponding simulated trajectories.

3.1. S1 specificity of three subtilases

The association constants for binding of engineered canonical protein inhibitors to subtilisin Carlsberg have been determined experimentally [43–45], and include 20 different P1 variants of OMTKY3 and 8 for eglin c. These studies show that the secondary interactions are slightly different as judged by the association constant for the P1 Gly variant, but with very similar $\Delta\Delta G(\mathrm{Gly} \rightarrow \mathrm{X})$ values. Similar $\Delta\Delta G(\mathrm{Gly} \rightarrow \mathrm{X})$ values means that the P1–S1 interaction for a given amino acid at P1 is nearly identical regardless of the protein inhibitor, and that the interaction free energies relative to a reference state (e.g., Gly) can be taken as a measure of the proteinase specificity. Amino acids that give a negative $\Delta\Delta G(\mathrm{Gly} \rightarrow \mathrm{X})$ are considered to have a favorable effect on the association energy, whereas the ones providing a positive $\Delta\Delta G(\mathrm{Gly} \rightarrow \mathrm{X})$ have a deleterious effect.

In order to investigate the specificity profile for subtilisin Carlsberg free energy calculations have been carried out, and average ligand (P1) surrounding interaction free energies for eglin c in solution and in the complex are presented in Table 2. The corresponding calculated binding free energies are also presented in Table 2 with available experimental free energies. In order to find the optimal nonpolar coefficient, we have minimized the difference in the calculated and experimental binding free energies relative to P1 Gly. This procedure provides us with a nonpolar coefficient of 0.58, which compares very well with values obtained previously [41]. It should be noted that in this procedure we have included both neutral and charged interaction energies for acidic P1 variants, which is discussed in more detail below. Our calculations show that the relative binding free energies varies from −4.9 to 3.0 kcal/mol, thus providing a theoretical span of approximately 7.9 kcal/mol excluding P1 Pro. This compares well with the variation observed experimentally for accommodation of OMTKY3 P1 variants to subtilisin Carlsberg [43], which is, again excluding Pro, 6 kcal/mol. It should also be noted that for the theoretical specificity span only the most favorable state for the ionic P1 variants are included. When P1 variants of OMTKY3 binds to subtilisin Carlsberg only 5 have a deleterious effect on the association energy [43], and the remaining 14 (excluding Pro) variants are within 3.6 kcal/mol of P1 Gly. The calculated $\Delta\Delta G(\mathrm{Gly} \rightarrow \mathrm{X})$ values presented here show that two amino acids have a deleterious effect on the association energy, namely Arg and Asp. Those that have a favorable interaction free energy relative to P1 Gly are all within 4.9 kcal/mol of the reference state.

According to our calculations the most preferred amino acid at the P1 position of eglin c bound to subtilisin Carlsberg is Tyr. The side chain of this residue is nicely stacked into the S1-site forming both hydrophobic interactions as well as hydrogen bonds through its hydroxyl group (Fig. 2). The P1–S1 interaction free energy for two of the P1 variants (Ile and Leu) is overestimated compared to their experimental values by approximately 3.4–4.3 kcal/mol, but the relative preference for Leu over Ile (1.5 kcal/mol) agrees with the experimental difference (2.4 kcal/mol). Opposed to this, the LIE calculations underestimate binding of the basic P1 Arg variant by 3.6 kcal/mol relative to P1 Gly. Subtilisin Carlsberg appears to accommodate basic side chains in their deprotonated form, as the calculations with a charged side chain is energetically

Table 2
P1-surrounding interaction energies and corresponding binding free energies (kcal/mol) for accommodation of P1 variants of eglin c to subtilisin Carlsberg

P1 variant	Eglin c–subtilisin Carlsberg		Eglin c		$\beta\Delta\langle V_{l-s}^{elec}\rangle$	$\alpha\Delta\langle V_{l-s}^{vdw}\rangle$	$\Delta\Delta G_{bind}^{exp}$ [a]	$\Delta\Delta G_{bind}^{LIE}$ [b]
	$\langle V_{l-s}^{elec}\rangle$	$\langle V_{l-s}^{vdw}\rangle$	$\langle V_{l-s}^{elec}\rangle$	$\langle V_{l-s}^{vdw}\rangle$				
Gly	−59.2	−8.2	−60.9	−3.7	0.7	−2.6	0.0	0.0
Leu	−59.3	−23.1	−62.0	−11.4	1.2	−6.8	−0.4	−3.8
Met	−62.1	−23.3	−64.9	−11.5	1.2	−6.9	nd	−3.8
Val	−59.3	−18.5	−61.0	−9.7	0.7	−5.1	nd	−2.5
Gln	−77.9	−23.2	−88.0	−8.9	4.3	−8.3	nd	−2.1
Phe	−62.6	−26.3	−66.7	−13.7	1.8	−7.4	nd	−3.7
Asn	−75.1	−18.8	−85.3	−7.2	4.4	−6.8	nd	−0.5
His[c,d]	−90.0	−21.7	−119.5	−9.3	14.8	−7.2	nd	11.4
His0[d]	−76.3	−22.8	−83.3	−10.8	3.0	−7.0	nd	−2.1
Cys	−60.1	−17.5	−64.9	−8.0	2.1	−5.5	nd	−1.6
Ala	−57.7	−13.5	−61.7	−6.1	1.7	−4.3	nd	−0.7
Tyr	−74.8	−25.8	−76.3	−13.2	0.6	−7.4	nd	−4.9
Ser	−76.0	−12.5	−75.0	−4.7	−0.4	−4.6	−1.6	−3.0
Thr	−69.9	−17.1	−72.2	−7.0	0.9	−5.9	nd	−3.1
Ile	−55.5	−22.1	−60.8	−11.6	2.0	−6.1	2.0	−2.3
Trp	−71.6	−30.8	−75.1	−17.0	1.5	−8.1	nd	−4.7
Asp[c]	−242.1	−7.3	−247.4	−0.1	2.7	−4.2	3.3	0.6
Asp0[d]	−72.4	−17.2	−80.5	−6.3	3.0	−6.4	3.3	4.3
Glu[c]	−241.6	−13.3	−247.9	−1.3	3.2	−7.0	−0.1	−1.7
Glu0[d]	−75.8	−20.4	−85.4	−7.8	3.6	−7.4	−0.1	3.2
Arg[c]	−101.0	−24.3	−126.0	−12.0	12.5	−7.2	−0.6	6.9
Arg0[d]	−101.5	−23.9	−110.7	−9.2	4.0	−8.6	−0.6	3.0
Lys[c]	−107.7	−21.6	−136.5	−8.5	14.4	−7.6	−1.5	8.3
Lys0[d]	−72.9	−24.2	−77.9	−10.1	2.2	−8.2	−1.5	−1.0

[a] Experimental data obtained from Refs. [44,45], while nd means not determined.
[b] Convergence errors in the calculated binding free energies are between 0.3 and 0.8 kcal/mol.
[c] The contribution from long-range electrostatics has been added, and ranges is −0.4, −0.3, −0.3, 0.3 and 0.3 kcal/mol for P1 Lys, His, Arg, Asp and Glu, respectively.
[d] 0 indicates neutral side chain. The free energy required to protonate/deprotonate the side chain has been added according to $\Delta\Delta G_{bind}^{pK_a} = 1.35|pH - pK_a|$ where pH is 8.3 (corresponding to the pH used in the association measurements) and pK_a is 12.5, 10.7, 6.6, 4.5 and 4.0 for Arg (5.7 kcal/mol), Lys (3.2 kcal/mol), His (2.3 kcal/mol), Glu (5.1 kcal/mol) and Asp (5.8 kcal/mol), respectively.

very unfavorable. In fact, there is a difference of 3.9–13.5 kcal/mol between the calculations based on the deprotonated form (P1 Arg0, His0 and Lys0 in Table 2) and those using a charged description (P1 Arg, His and Lys in Table 2) of the basic side chains. Thus, the calculations clearly show that all the basic amino acids are accommodated in neutral form (at pH 8.3), and not as positively charged variants to the S1-site of subtilisin Carlsberg. It should be noted that the binding free energies for neutral Arg and Lys and charged His have been corrected with the free energy required to deprotonate/protonate the side chain when it is free in solution (eglin c). Since the association energy measurements were determined at pH 8.3, the Arg and Lys side chain will be charged in eglin c while His will be neutral, and this contribution is given by $\Delta\Delta G_{bind}^{pK_a} = 1.35|pH - pK_a|$ where pH is 8.3 and pK_a is that of the free amino acid.

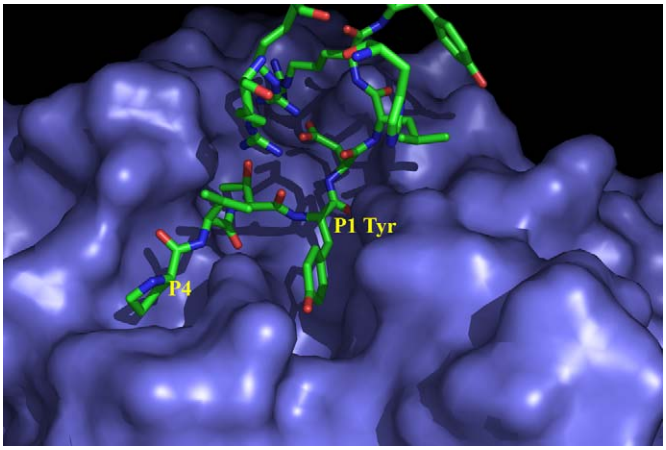

Fig. 2. Illustration of accommodation of eglin c P1 Tyr to the active site of subtilisin Carlsberg as observed in the MD simulations. Only the molecular surface of subtilisin Carlsberg is shown for clarity. The binding loop of eglin c is shown in sticks along with important residues in the vicinity of the binding site.

When it comes to the acidic variants Asp and Glu, the LIE calculations estimate a binding free energy difference of 0.6 and 4.3 kcal/mol for accommodation of charged and neutral Asp, while the corresponding values for Glu are −1.7 and 3.2 kcal/mol. Experimental binding free energies are available for P1 Asp and Glu and are 3.3 and −0.1 kcal/mol relative to P1 Gly, respectively. There is thus a preference for the charged form of both the acidic amino acids according to the calculations. When the calculations are carried out, one must assign a protonation state that is fixed during the simulation. If the pK_a value of the amino acid of interest is close to the pH where the association energy measurements are carried out, the experimental binding free energy will reflect contributions from two states, protonated and deprotonated. Conventional

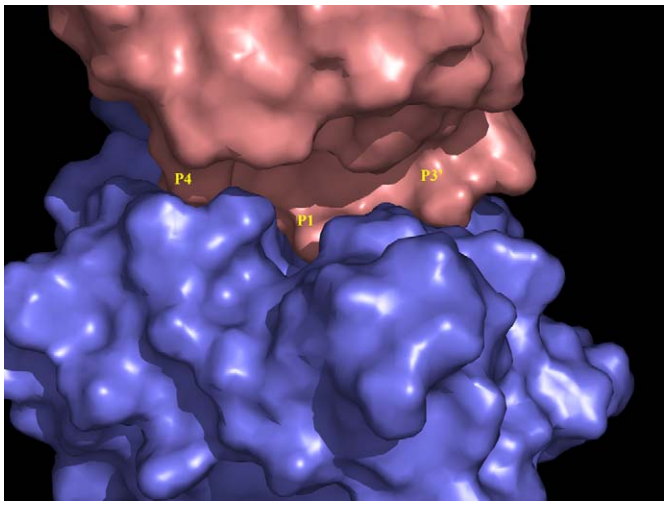

Fig. 3. View of the protein–protein interface between eglin c and subtilisin Carlsberg illustrating the solvent exposure of the P1–S1 interaction.

MD simulations do not allow such issues to be examined in a straightforward manner. Accommodation of both P1 Asp and Glu is overestimated when they are in the charged state, while underestimated in the protonated (neutral) state. If we assume that the pK_a is around pH 8.3 where the association measurements were conducted, it is reasonable to expect that the average of the binding free energies obtained with the charged and neutral topology of the P1 Asp and Glu is closer to the experimental free energies. For accommodation of P1 Asp the average of the relative binding free energies (P1 Asp and Asp0) is 2.5 kcal/mol, which is in much better agreement with the experimental value of 3.3 kcal/mol. Correspondingly, we obtain an average value of 0.8 kcal/mol for P1 Glu, which is again closer to the experimental value of −0.1 kcal/mol than that obtained with P1 Glu and Glu0. The assumption of a pK_a value of approximately 8.3 is supported by the findings of Qasim et al. [46], where the pK_a value of P1 Glu was found to be 8.7 in the complex between OMTKY3 and *Streptomyceus griseus* proteinase B. The calculated binding free energies involving the protonated acids have also been corrected with the free energy required to deprotonate the P1 side chain when eglin c is free in solution. Structural investigations also show that there is bulk solvent which have access to the P1 residue (Fig. 3). In fact Fig. 3 shows that there is a tunnel-like opening spanning across the protein–protein interface allowing solvent to interact with the P1 residue. Even though the P1 residues are not fully exposed to solvent in the complex significant stabilization of the charged variants is achieved by allowing water molecules to interact with especially the ionizable P1 variants.

Table 3
P1-surrounding interaction energies and corresponding binding free energies (kcal/mol) for binding of P1 variants of eglin c to thermitase

P1 variant	Eglin c–thermitase		Eglin c		$\beta\Delta\langle V_{l-s}^{elec}\rangle$	$\alpha\Delta\langle V_{l-s}^{vdw}\rangle$	$\Delta\Delta G_{bind}^{LIE}$ [a]
	$\langle V_{l-s}^{elec}\rangle$	$\langle V_{l-s}^{vdw}\rangle$	$\langle V_{l-s}^{elec}\rangle$	$\langle V_{l-s}^{vdw}\rangle$			
Gly	−60.7	−9.0	−62.1	−3.7	0.6	−3.1	0.0
Leu	−60.5	−22.6	−61.9	−11.5	0.6	−6.5	−3.4
Met	−64.5	−23.1	−64.4	−11.7	0.0	−6.7	−4.2
Val	−59.3	−20.1	−61.1	−9.6	0.8	−6.1	−2.9
Gln	−76.9	−22.8	−87.6	−8.6	4.6	−8.3	−1.2
Phe	−59.5	−27.2	−66.3	−13.9	2.9	−7.8	−2.3
Asn	−79.2	−18.2	−86.5	−7.0	3.1	−6.5	−0.9
His[b,c]	−52.4	−21.3	−78.5	−9.3	13.1	−7.0	10.5
His0[c]	−74.4	−25.3	−82.8	−10.6	3.6	−8.6	−2.5
Cys	−61.3	−16.3	−65.1	−7.7	1.6	−5.0	−0.9
Ala	−59.7	−12.6	−61.6	−5.7	0.8	−4.0	−0.7
Tyr	−67.3	−28.3	−76.1	−13.2	3.3	−8.8	−3.1
Ser	−71.9	−13.4	−75.2	−4.7	1.2	−5.1	−1.4
Thr	−71.8	−17.4	−73.6	−7.1	0.7	−6.0	−2.9
Ile	−57.7	−23.7	−60.8	−11.2	1.1	−7.3	−3.7
Trp	−68.7	−32.7	−75.2	−16.7	2.8	−9.3	−4.0
Asp[b]	−284.4	−7.7	−285.8	0.0	0.7	−4.5	−1.2
Asp0[c]	−77.1	−15.7	−80.2	−6.3	1.1	−5.5	4.0
Glu[b]	−282.6	−13.8	−288.7	−1.5	3.0	−7.2	−1.4
Glu0[c]	−80.1	−21.0	−82.8	−8.2	1.0	−7.5	1.1
Arg[b]	−63.2	−25.5	−84.9	−11.6	10.9	−8.1	4.9
Arg0[c]	−107.1	−24.0	−111.6	−8.3	1.9	−9.2	1.0
Lys[b]	−63.0	−21.2	−94.2	−8.9	15.6	−7.2	10.8
Lys0[c]	−76.0	−23.3	−78.7	−10.4	1.2	−7.5	−0.7

[a] Covergence errors in the calculated binding free energies are between 0.3 and 0.8 kcal/mol.
[b] The contribution from long-range electrostatics has been added, and ranges is −0.1, −0.3, −0.3, 0.1 and 0.2 kcal/mol for P1 Lys, His, Arg, Asp and Glu, respectively.
[c] 0 indicates neutral side chain. The free energy required to protonate/deprotonate the side chain has been added according to $\Delta\Delta G_{bind}^{pK_a} = 1.35|pH - pK_a|$ where pH is 8.3 (corresponding to the pH used in the association measurements) and pK_a is 12.5, 10.7, 6.6, 4.5 and 4.0 for Arg (5.7 kcal/mol), Lys (3.2 kcal/mol), His (2.3 kcal/mol), Glu (5.1 kcal/mol) and Asp (5.8 kcal/mol), respectively.

The other two proteinases that were included in this study are thermitase and proteinase K, but much less are known about these two enzymes as compared to subtilisin Carlsberg. Thermitase consists of 279 residues while proteinase K has 277 residues, and the sequence identity between thermitase and proteinase K is 44%. Proteinase K has two disulfide bridges and a free cysteine residue while thermitase is characterized by a free cysteine residue near the active site. In general, thermitase has a preference for large uncharged residues in S1 and S4 (secondary binding pocket) while proteinase K is considered as relatively unspecific.

The average P1-surrounding interaction energies for binding of eglin c to thermitase and proteinase K are presented in Tables 3 and 4, respectively, and the calculated binding free energy relative to Gly is also given for each P1 variant. For both thermitase and proteinase K the electrostatic contribution relative to P1 Gly opposes binding, as only 1 and 3 variants have a favorable electrostatic contribution (relative to Gly), respectively. Furthermore, there are not any P1 side chains that have a more favorable electrostatic interaction in the complex compared to the unbound solvated eglin c, confirming that the nature of the S1-site is predominantly hydrophobic for both enzymes.

3.2. Comparison of the predicted S1 specificity

Table 5 shows a comparison of the calculated binding free energies relative to P1 Gly for accommodation of P1 variants of eglin c to subtilisin Carlsberg, thermitase and proteinase K. Very similar specificity profiles are found for all three enzymes, which is reflected in the correlation coefficients (r) calculated from the data in Table 5; subtilisin Carlsberg/thermitase 0.95, subtilisin Carlsberg/proteinase K 0.94 and thermitase/proteinase K 0.96. Thus, primary (S1) specificity of the three enzymes is similar and the introduction of different side chains at the P1-site results in equivalent energetic effects.

It can also be observed from Table 5 that basic amino acids are strongly disfavored when accommodated in their charged states in all three enzymes, and that acidic variants can bind in their charged form. The positively charged residue His is the least preferred in subtilisin Carlsberg while Lys is the least preferred for thermitase and proteinase K. The P1 side chain in all standard mechanism canonical protein inhibitors of serine proteinases is highly exposed to solvent [1]. Therefore, when the P1 side chain is charged, one can assume that the charge resides in a high dielectric medium. Upon transfer to the hydrophobic S1 cavity, it is forced to reside in a much lower dielectric environment. Transfers of a charge from high to low dielectric are strongly disfavored. This points in the direction that ionizable amino acids are accommodated as neutral variants, but the data presented here (Table 5) shows that the acidic residues are able to bind in their charged form. A closer inspection of the MD simulations reveals that the eglin c has two arginine residues and a histidine residue that is placed in

Table 4
P1–surrounding interaction energies and corresponding binding free energies (kcal/mol) for binding of P1 variants of eglin c to proteinase K

P1 variant	Eglin c–proteinase K		Eglin c		$\beta\Delta\langle V_{l-s}^{elec}\rangle$	$\alpha\Delta\langle V_{l-s}^{vdw}\rangle$	$\Delta\Delta G_{bind}^{LIE}$ [a]
	$\langle V_{l-s}^{elec}\rangle$	$\langle V_{l-s}^{vdw}\rangle$	$\langle V_{l-s}^{elec}\rangle$	$\langle V_{l-s}^{vdw}\rangle$			
Gly	−57.4	−9.8	−60.9	−3.9	1.5	−3.4	0.0
Leu	−56.9	−23.6	−62.0	−11.7	2.2	−6.9	−2.3
Met	−57.9	−23.3	−64.5	−11.9	2.8	−6.7	−1.4
Val	−56.0	−19.7	−60.1	−10.5	1.8	−5.4	−1.4
Gln	−73.7	−22.7	−87.7	−8.9	6.0	−8.1	0.6
Phe	−61.6	−27.1	−66.6	−13.9	2.2	−7.7	−3.0
Asn	−73.6	−19.8	−86.6	−7.2	5.6	−7.4	0.8
His[b,c]	−86.1	−22.4	−119.3	−9.9	16.6	−7.3	14.1
His0[c]	−70.3	−24.9	−83.2	−10.6	5.5	−8.3	−0.1
Cys	−58.5	−18.5	−64.9	−8.1	2.8	−6.1	−1.0
Ala	−57.6	−14.5	−61.7	−5.7	1.8	−5.1	−1.2
Tyr	−64.6	−29.4	−75.7	−13.5	4.1	−9.3	−2.4
Ser	−70.8	−15.0	−74.0	−5.1	1.2	−5.8	−2.3
Thr	−67.8	−19.6	−71.1	−7.4	1.2	−7.1	−3.4
Ile	−58.3	−22.4	−60.4	−11.8	0.8	−6.2	−2.9
Trp	−70.2	−33.5	−75.8	−17.0	2.4	−9.6	−4.4
Asp[b]	−231.2	−14.2	−250.2	−0.2	9.5	−8.2	4.0
Asp0[c]	−67.4	−19.6	−81.2	−5.9	5.1	−8.0	5.5
Glu[b]	−242.5	−14.3	−248.0	−1.1	2.8	−7.7	−2.4
Glu0[c]	−75.9	−20.1	−80.3	−9.1	1.6	−6.4	2.7
Arg[b]	−91.1	−29.5	−126.7	−10.5	17.8	−11.1	13.7
Arg0[c]	−101.6	−26.7	−111.4	−8.4	4.2	−10.7	2.7
Lys[b]	−92.0	−21.8	−136.9	−8.6	22.5	−7.7	17.3
Lys0[c]	−67.4	−25.5	−78.0	−10.1	4.6	−9.0	1.6

[a] Convergence errors in the calculated binding free energies are around 0.3–0.8 kcal/mol.
[b] The contribution from long-range electrostatics has been added, and ranges is −0.1, −0.1, 0.0, 0.0 and 0.1 kcal/mol for P1 Lys, His, Arg, Asp and Glu, respectively.
[c] 0 indicates neutral side chain. The free energy required to protonate/deprotonate the side chain has been added according to $\Delta\Delta G_{bind}^{pK_a} = 1.35|pH - pK_a|$ where pH is 8.3 (corresponding to the pH used in the association measurements) and pK_a is 12.5, 10.7, 6.6, 4.5 and 4.0 for Arg (5.7 kcal/mol), Lys (3.2 kcal/mol), His (2.3 kcal/mol), Glu (5.1 kcal/mol) and Asp (5.8 kcal/mol), respectively.

Table 5
Comparison of calculated relative binding free energies $\Delta\Delta G(\text{Gly} \rightarrow \text{X})$ (kcal/mol) for binding of P1 variants of eglin c to subtilisin Carlsberg, thermitase and proteinase K

P1 variant	Subtilisin Carlsberg	Thermitase	Proteinase K
Gly	0.0	0.0	0.0
Leu	−3.8	−3.4	−2.3
Met	−3.8	−4.2	−1.4
Val	−2.5	−2.9	−1.4
Gln	−2.1	−1.2	0.6
Phe	−3.7	−2.3	−3.0
Asn	−0.5	−0.9	0.8
His	11.4	10.5	14.1
His0[a]	−2.1	−2.5	−0.1
Cys	−1.6	−0.9	−1.0
Ala	−0.7	−0.7	−1.2
Tyr	−4.9	−3.1	−2.4
Ser	−3.0	−1.4	−2.3
Thr	−3.1	−2.9	−3.4
Ile	−2.3	−3.7	−2.9
Trp	−4.7	−4.0	−4.4
Asp	0.6	−1.2	4.0
Asp0[a]	4.3	4.0	5.5
Glu	−1.7	−1.4	−2.4
Glu0[a]	3.2	1.1	2.7
Arg	6.9	4.9	13.7
Arg0[a]	3.0	1.0	2.7
Lys	8.3	10.8	17.3
Lys0[a]	−1.0	−0.7	1.6

[a] Indicates binding of uncharged amino acid.

close vicinity to the S1-site upon complex formation, and these may provide a stabilization of the negatively charged P1 variants. These three residues also give an additional destabilization of the basic Lys, Arg and His side chains, which in part explains the specificity preference when accommodating ionizable variants.

The landscape of the S1-site of subtilases is generally more flat compared to the other serine proteinase, for example from the chymotrypsin-like family. Comparison of the S1-site of subtilisin Carlsberg, thermitase and proteinase K is given in Fig. 4. Fig. 4 shows that the S1–S4 binding sites are rather flat, and this may be an explanation for the broad specificity observed for these enzymes. Furthermore, very similar curvature is observed among the three enzymes studied here, which adds support to similar energetic effects observed upon mutations of the P1 residue of eglin c.

In summary, we have used free energy calculations to investigate the effect of mutating the P1 residue in eglin c when bound to subtilisin Carlsberg, thermitase and proteinase K. Previous application of the LIE method to study the effect of point mutations in complexes between chymotrypsin and elastase with OMTKY3 and trypsin with BPTI provided a mean α of 0.58. We obtain exactly the same value when using the eight experimental association constants for subtilisin Carlsberg to optimize the α, indicating that this value is generally applicable to proteinase–protein inhibitor complexes. However, more studies of other protein–protein interfaces are needed in order to fully resolve the possible system dependencies of the α parameter. It might be the case

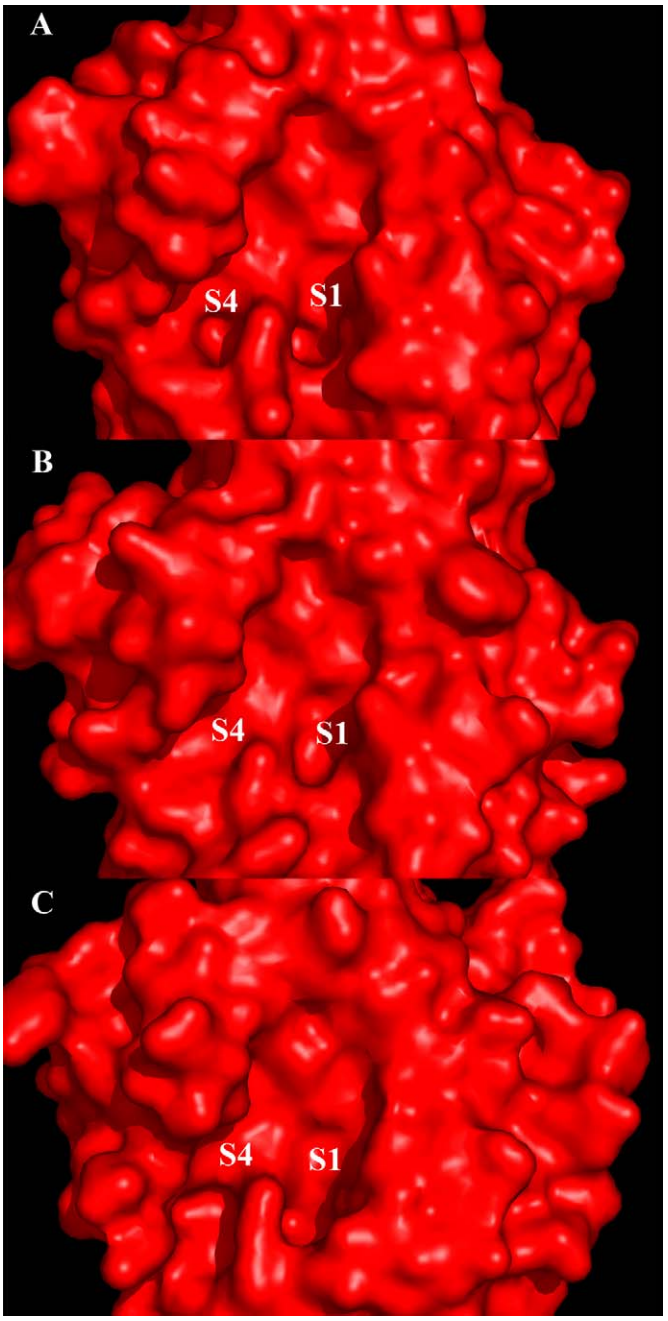

Fig. 4. Comparison of the surface at the S1 and S4 binding sites of subtilisin Carlsberg (A), thermitase (B) and proteinase K (C).

that this parameter will attain different values depending on the protein systems studied.

One important question related to accommodation of ionic P1 variants is the pK_a changes that might accompany binding processes. The present calculations show that the basic side chain Arg, Lys and His are preferred as neutral (deprotonated) when bound to all three subtilases investigated. In contrast, all acidic variants are most preferred as charged side chains for all three enzymes. However, due to the limited amount of experimental data, it is difficult to fully address the issues of pK_a changes in the P1 side chain upon complex formation. Nonetheless, the LIE calculations presented here show that the energetic effect of point

mutations of the P1 residue is very similar in the three subtilases investigated. To fully address the generality and suitability of the LIE method to probe the effect of point mutations at protein–protein interfaces more interfaces must be examined.

Acknowledgements

Financial support from the Research Council of Norway is gratefully appreciated. The Norwegian Structural Biology Centre (NorStruct) is supported by the Functional Genomics Program (FUGE) of the Research Council of Norway.

References

[1] W. Bode, R. Huber, Natural protein proteinase inhibitors and their interaction with proteinases, Eur. J. Biochem. 204 (1992) 433–451.
[2] M. Laskowski, I. Kato, Protein inhibitors of proteinases, Annu. Rev. Biochem. 49 (1980) 593–626.
[3] S. Jones, J. Thornton, Principles of protein–protein interactions, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 13–20.
[4] J. Otlewski, D. Krowarsch, W. Apostoluk, Protein inhibitors of serine proteinases, Acta Biochim. Pol. 46 (1999) 531–565.
[5] I. Schechter, A. Berger, On size of active site in proteases I. Papain, Biochem. Biophys. Res. Commun. 27 (1967) 157–162.
[6] M. Laskowski, I. Apostol, W. Ardelt, J. Cook, A. Giletto, C. Kelly, W. Lu, S. Park, M. Qasim, H. Whatley, A. Wieczorek, R. Wynn, Amino acid sequences of ovomucoid 3rd domain from 25 additional species of birds, J. Prot. Chem. 9 (1990) 715–725.
[7] U. Seemuller, M. Meier, K. Ohlsson, H. Muller, H. Fritz, Isolation and characterization of a low molecular weight inhibitor (of chymotrypsin and human granulocytic elastase and cathepsin G) from leeches, H.-S. Z. Physiol. Chem. 358 (1977) 1105–1117.
[8] U. Seemuller, M. Eulitz, H. Fritz, A. Strobl, Structure of the elastase cathepsin G inhibitor of the leech *Hirudo medicinalis*, H.-S. Z. Physiol. Chem. 361 (1980) 1841–1846.
[9] C. Betzel, Z. Dauter, N. Genov, V. Lamzin, J. Navaza, H. Schnebli, M. Visanji, K. Wilson, Structure of the proteinase inhibitor eglin c with hydrolyzed reactive center at 2.0 angstrom resolution, FEBS Lett. 317 (1993) 185–188.
[10] S. Hyberts, M. Goldberg, T. Havel, G. Wagner, The solution structure of eglin c based on measurements of many NOEs and coupling coupling constants and its comparison with X-ray structures, Protein Sci. 1 (1992) 736–751.
[11] W. Bode, E. Papamokos, D. Musil, The high-resolution X-ray crystal structure of the complex formed between subtilisin Carlsberg and eglin c, an elastase inhibitor from the leech *Hirudo medicinalis*. Structural analysis, subtilisin structure and interface geometry, Eur. J. Biochem. 166 (1987) 673–692.
[12] F. Frigerio, A. Coda, L. Pugliese, C. Lionetti, E. Menegatti, G. Amiconi, H.P. Schnebli, P. Ascenzi, M. Bolognesi, Crystal and molecular structure of the bovine alpha-chymotrypsin–eglin c complex at 2.0 A resolution, J. Mol. Biol. 225 (1992) 107–123.
[13] R.J. Siezen, J.A. Leunissen, Subtilases: the superfamily of subtilisin-like serine proteases, Protein Sci. 6 (1997) 501–523.
[14] J. Arnorsdottir, R.B. Smaradottir, O.T. Magnusson, S.H. Thorbjarnardottir, G. Eggertsson, M.M. Kristjansson, Characterization of a cloned subtilisin-like serine proteinase from a psychrotrophic Vibrio species, Eur. J. Biochem. 269 (2002) 5536–5546.
[15] S. Davail, G. Feller, E. Narinx, C. Gerday, Cold adaptation of proteins. Purification, characterization, and sequence of the heat-labile subtilisin from the antarctic psychrophile Bacillus TA41, J. Biol. Chem. 269 (1994) 17448–17453.
[16] J.A. Wells, E. Ferrari, D.J. Henner, D.A. Estell, E.Y. Chen, Cloning, sequencing, and secretion of *Bacillus amyloliquefaciens* subtilisin in *Bacillus subtilis*, Nucleic Acids Res. 11 (1983) 7911–7925.
[17] G. Hausdorf, K. Kruger, W.E. Hohne, Thermitase, a thermostable serine protease from *Thermoactinomyces vulgaris*. Classification as a subtilisin-type protease, Int. J. Pept. Protein Res. 15 (1980) 420–429.
[18] K. Peek, R.M. Daniel, C. Monk, L. Parker, T. Coolbear, Purification and characterization of a thermostable proteinase isolated from *Thermus* sp. strain Rt41A, Eur. J. Biochem. 207 (1992) 1035–1044.
[19] M.W. Pantoliano, M. Whitlow, J.F. Wood, S.W. Dodd, K.D. Hardman, M.L. Rollence, P.N. Bryan, Large increases in general stability for subtilisin BPN′ through incremental changes in the free energy of unfolding, Biochemistry 28 (1989) 7205–7213.
[20] P. Volkl, P. Markiewicz, K.O. Stetter, J.H. Miller, The sequence of a subtilisin-type protease (aerolysin) from the hyperthermophilic archaeum *Pyrobaculum aerophilum* reveals sites important to thermostability, Protein Sci. 3 (1994) 1329–1340.
[21] E. Narinx, E. Baise, C. Gerday, Subtilisin from psychrophilic antarctic bacteria: characterization and site-directed mutagenesis of residues possibly involved in the adaptation to cold, Protein Eng. 10 (1997) 1271–1279.
[22] P.L. Wintrode, K. Miyazaki, F.H. Arnold, Cold adaptation of a mesophilic subtilisin-like protease by laboratory evolution, J. Biol. Chem. 275 (2000) 31635–31640.
[23] R. Kleine, Properties of thermitase, a thermostable serine protease from *Thermoactinomyces vulgaris*, Acta Biol. Med. Ger. 41 (1982) 89–102.
[24] G. Voordouw, G.M. Gaucher, R.S. Roche, Physiochemical properties of thermomycolase, the thermostable, extracellular, serine protease of the fungus *Malbranchea pulchella*, Can. J. Biochem. 52 (1974) 981–990.
[25] D.N. Georgieva, S. Stoeva, W. Voelter, N. Genov, C. Betzel, Differences in the specificities of the highly alkalophilic proteinases Savinase and Esperase imposed by changes in the rigidity and geometry of the substrate binding sites, Arch. Biochem. Biophys. 387 (2001) 197–201.
[26] K.J. Stevenson, G.M. Gaucher, The substrate specificity of thermomycolase, an extracellular serine proteinase from the thermophilic fungus *Malbranchea pulchella* var. sulfurea, Biochem. J. 151 (1975) 527–542.
[27] J. Åqvist, C. Medina, J. Samulesson, New method for predicting binding-affinity in computer-aided drug design, Protein Eng. 7 (1994) 385–391.
[28] P. Gros, C. Betzel, Z. Dauter, K.S. Wilson, W.G.J. Hol, Molecular-dynamics refinement of a thermitase eglin-c complex at 1.98 Å resolution and comparison of 2 crystal forms that differ in calcium content, J. Mol. Biol. 210 (1989) 347–367.
[29] T.A. Jones, J.Y. Zou, S.W. Cowan, M. Kjeldgaard, Improved methods for building protein models in electron-density maps and the location of errors in these models, Acta Cryst. Sec. A 47 (1991) 110–119.
[30] C. Betzel, S. Gourinath, P. Kumar, P. Kaur, M. Perbandt, S. Eschenburg, T.P. Singh, Structure of a serine protease proteinase K from *Tritirachium album* limber at 0.98 A resolution, Biochemistry 40 (2001) 3080–3088.
[31] J. Marelius, K. Kolmodin, I. Feierberg, J. Aqvist, Q: a molecular dynamics program for free energy calculations and empirical valence bond simulations in biomolecular systems, J. Mol. Graph. Model. 16 (1998) 213–225, 261.
[32] W. Cornell, P. Cieplack, C. Bayly, I. Gould, K. Merz, D. Ferguson, D. Spellmeyer, T. Fox, J. Caldwell, P. Kollman, A 2nd generation force field for the simulation of proteins, nucleic acids, and organic molecules, J. Am. Chem. Soc. 117 (1995) 5179–5197.
[33] F.S. Lee, A. Warshel, A local reation field method for fast evaluation of long-range electrostatic interactions in molecular simulations, J. Chem. Phys. 97 (1992) 3100–3107.
[34] W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, M. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[35] B. Brandsdal, J. Åqvist, A. Smalås, Computational analysis of binding of P1 variants to trypsin, Protein Sci. 10 (2001) 1584–1595.
[36] J. Åqvist, Calculation of absolute binding free energies for charged ligands and effects of long-range electrostatic interactions, J. Comput. Chem. 17 (1996) 1587–1597.

[37] B. Brandsdal, F. Osterberg, M. Almlöf, I. Feierberg, V. Luzhkov, J. Åqvist, Free energy calculations and ligand binding, Adv. Protein Chem. 66 (2003) 123–158.
[38] J. Åqvist, V.B. Luzhkov, B.O. Brandsdal, Ligand binding affinities from MD simulations, Acc. Chem. Res. 35 (2002) 358–365.
[39] M. Almlöf, B.O. Brandsdal, J. Åqvist, Binding affinity prediction with different force fields: examination of the linear interaction energy method, J. Comput. Chem. 25 (2004) 1242–1254.
[40] T. Hansson, J. Marelius, J. Åqvist, Ligand binding affinity prediction by linear interaction energy methods, J. Comput. Aided Mol. Des. 12 (1998) 27–35.
[41] M. Almlöf, J. Åqvist, A.O. Smalås, B.O. Brandsdal, Probing the effect of point mutations at protein–protein interfaces with free energy calculations, Biophys. J. 90 (2006) 1–10.
[42] F.S. Lee, Z.T. Chu, M.B. Bolger, A. Warshel, Calculations of antibody-antigen interactions: microscopic and semi-microscopic evaluation of the free energies of binding of phosphorylcholine analogs to McPC603, Protein Eng. 5 (1992) 215–228.
[43] W. Lu, I. Apostol, M. Qasim, N. Warne, R. Wynn, W. Zhang, S. Anderson, Y. Chiang, E. Ogin, I. Rothberg, K. Ryan, M. Laskowski, Binding of amino acid side chains to S1 cavities of serine proteinases, J. Mol. Biol. 266 (1997) 441–461.
[44] M. Qasim, P. Ganz, C. Saunders, K. Bateman, M. James, M. Laskowski, Interscaffolding additivity. Association of P1 variants of eglin c and of turkey ovomucoid third domain with serine proteinases, Biochemistry 36 (1997) 1598–1607.
[45] M.A. Qasim, S.M. Lu, J. Ding, K.S. Bateman, M.N. James, S. Anderson, J. Song, J.L. Markley, P.J. Ganz, C.W. Saunders, M. Laskowski Jr., Thermodynamic criterion for the conformation of P1 residues of substrates and of inhibitors in complexes with serine proteinases, Biochemistry 38 (1999) 7142–7150.
[46] M.A. Qasim, M.R. Ranjbar, R. Wynn, S. Anderson, M.J. Laskowski, Ionizable P1 residues in serine proteinase inhibitors undergo large pK shifts on complex formation, J. Biol. Chem. 270 (1995) 27419–27422.