Challenges in pK_a Predictions for Proteins: The case of Asp213 in Human Proteinase 3[†]

Eric Hajjar,[†,§] Annick Dejaegere,[‡] and Nathalie Reuter*[,§,||]

Department of Physics, University of Cagliari (CA), Italy, Biocomputing Group, IGBMC, Université de Strasbourg, Illkirch, France, and Computational Biology Unit, BCCS, Department of Molecular Biology, University of Bergen, Norway

Received: March 31, 2009; Revised Manuscript Received: September 12, 2009

Knowledge of the protonation states of the ionizable residues in an enzyme is a prerequisite to an accurate description of its structure and mechanism. In practice, the use of the inappropriate protonation state for an amino acid in a molecular modeling computation (e.g., molecular dynamics simulation) is likely to lead to unrealistic results. Although methods using solvers of the linearized Poisson−Boltzmann equation have proven to yield accurate pK_a predictions, they bear a number of limitations. They are quite demanding in terms of computational power and are sensitive to representation of the charges and their position (force field and protein conformation). Moreover they depend on the choice of a dielectric constant for the protein interior. In this manuscript, we describe the difficulties met when trying to predict the protonation state of a buried amino acid, located in a protein for which very little biochemical data is available. Such a case is highly representative of the challenges faced in theoretical biology studies. Proteinase 3 (PR3) is an enzyme involved in proteolytic events associated with inflammation. It is a potential target in the development of new anti-inflammatory therapeutic strategies. We report the results of pK_a predictions of the aspartic acid 213 of PR3 with a FDPB solver. We probed the influence of the choice of the dielectric constant for the protein interior ε_p and the benefits of conformational sampling by molecular dynamics (MD) on the pK_a prediction of this carboxylate group. Using only the FDPB calculations, we could not conclude on the protonation state of Asp213. MD simulations confronted to knowledge of the ligand-binding and reaction mechanism led us to decide on a protonated form of this aspartic acid. We also demonstrate that the use of the wrong protonation state leads to an unreliable structural model for PR3. pK_a prediction with a fast empirical method yielded a pK_a of 8.4 for Asp213, which is in agreement with our choice of protonation state based on MD simulations.

Introduction

The importance of pH on the stability, ligand binding and catalytic activity of enzymes is well-known. The knowledge of the protonation states, and hence of the pK_a values, of the ionizable residues in an enzyme is thus a prerequisite to an accurate description of its structure and mechanism.[1−4] In practice, the use of the inappropriate protonation state for an amino acid in a molecular modeling computation (e.g., molecular dynamics simulation) is likely to lead to unrealistic results.[5] The establishment and use of reliable titration methods and protocols is thus essential in in silico studies of biomolecular systems aiming for example at describing ligand−receptor interactions, protein−protein complexes or designing drugs.

A number of methods to predict pK_a of proteins ionizable residues have been developed in the last decades (e.g., refs 6−13). Most of them are based on the fundamental concept that pK_a shifts in proteins result primarily from electrostatic interactions. The most popular methods use solvers of the linearized Poisson−Boltzmann equation (LPBE) to estimate the energetics of the ionization equilibria. In these models, the protein and solvent each have a distinct dielectric constant. The solvent surrounding the protein is a uniform medium with a high dielectric constant ($\varepsilon_{out} = 80$) while the protein is attributed a low dielectric constant ($\varepsilon_p = 4-20$). The charge distribution of the protein is usually represented by the partial charges of the force field, and the protein shape is defined by its tridimensional structure and by the radius of its solvent accessible atoms. The electrostatic potential is given everywhere by the Poisson−Boltzmann equation. Although no analytic solution exists for irregularly shaped systems, numerical solutions of the Poisson−Boltzmann (PB) equation can be obtained by finite difference (FD) (or boundary elements or finite elements) methods. FDPB methods have proven to yield accurate pK_a predictions, but they have several limitations; they are quite demanding in terms of computational power; they depend on the representation of the charges and their position (force field and protein conformation); and they are sensitive to the choice of the protein dielectric constant. Several studies have been dedicated to the influence of these parameters. The importance of taking the protein flexibility into account has been investigated for a number of proteins for which experimental pK_a values are available.[14−17] The use of molecular dynamics (MD) simulations to generate conformational ensembles has been shown to improve pK_a predictions.[14,15] Titration results obtained on single averaged structures from MD trajectories have been compared to results obtained from averaged titration curves calculated on several structures. The improvement provided by one strategy or the other differs for the different proteins studied,[14,15] so that no general protocol could be established. Moreover, Nielsen and

[†] Part of the "Walter Thiel Festschrift".

* To whom correspondence should be addressed. Address: Computational Biology Unit, BCCS, University of Bergen, Thormohlensgt 55, N-5008 Bergen, Norway. Tel: (47) 555 84040. Fax: (47) 555 84295. E-mail: nathalie.reuter@cbu.uib.no.

[†] University of Cagliari.

[‡] Université de Strasbourg.

[§] Computational Biology Unit, BCCS, University of Bergen.

[||] Department of Molecular Biology, University of Bergen.

10.1021/jp902930u CCC: $40.75
Published on Web 09/25/2009

McCammon[18] and Wlodek et al.[5] showed on selected examples that MD simulations and other forms of conformational averaging do not systematically improve predicted pK_a values. A major bias is that, since MD simulations usually require an initial choice of protonation states for titratable residues, the protein tends to relax and adapt to this initial protonation state. As a result, this initial protonation state often appears artificially favorable in titrations based on the MD trajectory.[5,14,18]

The use of a single or of several protein structures also influences the choice of the protein dielectric constant. When resorting to a single X-ray structure using no conformational sampling,[8,15,19] the choice of high values for ε_p (e.g., $\varepsilon_p = 20$) has been shown to yield more accurate results. For predictions based on conformational ensembles (average titration curves)[14,15] low values (ca. $\varepsilon_p = 4$) are recommended. It has also been described that pK_a values calculated on conformational ensembles are less sensitive to ε_p values.[14]

Methods using alternative approaches to FDPB solvers have also been reported; Generalized Born (GB) continuum solvent,[20,21] protein dipoles Langevin dipoles[11] (PDLD) model, or the fast empirical approach propKa[22,23] of Jensen and co-workers. In propKa, the pK_a shifts of amino acids depend on mainly three types of perturbations, desolvation, hydrogen bonds, and charge−charge interactions of buried atoms. PropKa has been tested against experimental pK_a values available for 44 proteins. Moreover pK_a values for five proteins used as benchmarks in previous studies have been calculated using propKa, and results of comparable accuracy to those of the Poisson−Boltzmann approaches were obtained.[22]

Experimental pK_a values are more readily accessible for solvent exposed amino acids possibly biasing the benchmarks studies toward accurate in silico predictions of surface residues. It has been suggested that conformational sampling affects mostly these amino acids.[18] However active site and buried residues constitute the vast majority of the functionally important amino acids in enzymes and are more affected by the protein environment. Their unusual titration curves are even used for protein functional site prediction.[24−26] They might require different computational strategies or titration protocols. Unfortunately less experimental data is available for these amino acids and the difficulties in setting up the right in silico titration strategy for them are poorly documented. In this manuscript we describe the difficulties met when trying to predict the protonation state of a buried amino acid, located in a protein for which very little biochemical data is available. Such a case is highly representative of the challenges faced in theoretical biology studies.

Proteinase 3 (PR3) is a serine protease of the neutrophils, the most abundant type of white blood cells and is involved in proteolytic events associated with inflammation. It is associated to several diseases such as emphysema, cystic fibrosis, rheumatoid arthritis, or vasculitis. It is a potential target in the development of new anti-inflammatory therapeutic strategies. The active site comprises the catalytic triad that consists of three amino acids, His57, Asp102, and Ser195 (Figure 1). The relative orientation of the histidine, serine, and aspartic acid is similar in all serine proteases and results in the formation of strong hydrogen bonds between histidine and serine on the one hand, and histidine and aspartic acid on the other hand. Upon substrate binding, the first step of the catalytic reaction is the attack of the catalytic serine on the carbonyl group of the cleavable amide bond and the transfer of the hydroxyl hydrogen of the serine to the histidine. This leads to the release of the C-terminal end of the substrate and the formation of a covalent intermediate, the

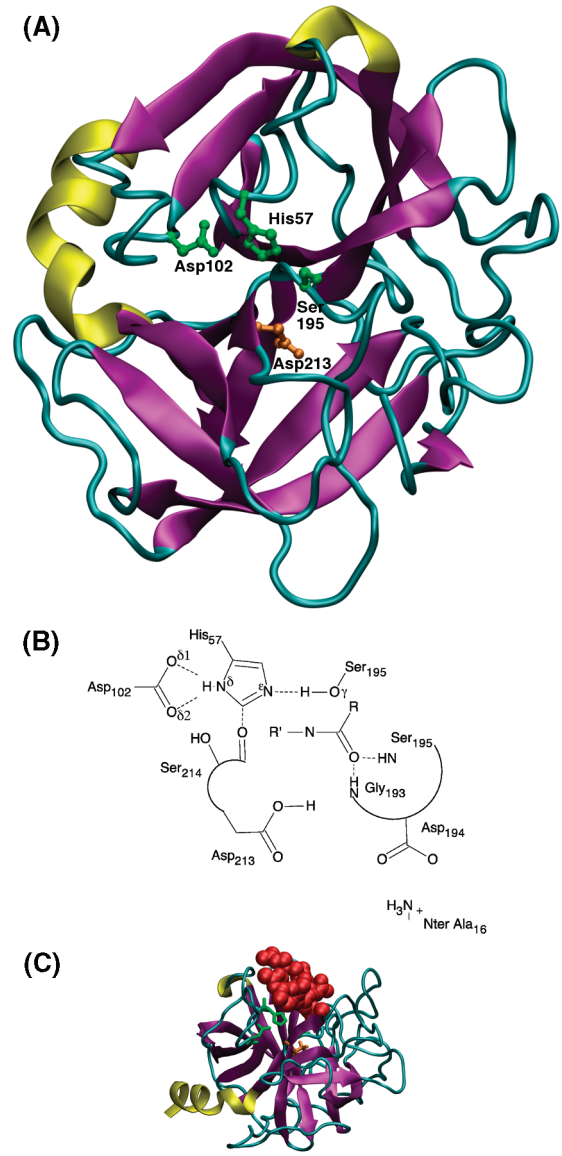

Figure 1. Stucture and active site of Proteinase 3. (A) Three-dimensional structure of PR3.[33] The secondary structure elements are represented by cartoons (purple arrows, extended strands; yellow, helices). The catalytic triad is represented in green balls and sticks and Asp213 in orange. (B) Two-dimensional representation of the catalytic triad (His57, Asp102, and Ser195), the oxyanion hole (NH of Gly193 and Ser195), Asp213 as well as Asp194 and the N-terminal group (Ile 16). Ile16 and Asp194 interact via a strong hydrogen bond ($NH_3^+ - {}^-OOC$) in the mature form of the enzyme. (C) Proteinase 3 bound to a cleavable heptapeptide (snapshot from the MD simulations of Proteinase 3 bound to VADVKDR).

acylenzyme, between the enzyme and the N-terminal part of the substrate. The substrates are positioned optimally in the active site thanks to a network of interactions that extends on both sides of the cleavable bond.[27] We and others have described these sites for PR3[28−32] and shown that some of them are significantly different from the ones of its closest homologue, the human neutrophil elastase (HNE); one striking difference between HNE and PR3 is the presence of the aspartic acid Asp213 at the bottom of the main binding site of PR3. While there is abundant structural information for HNE, there exist only one X-ray structure of Proteinase 3[33] (PDB ID: 1FUJ). The authors observed that the side-chain Asp213 points toward the core of the protein, in a position to act as a hydrogen bond donor to nearby residues. They therefore suggested that Asp213

might be protonated. However the extent to which a protein X-ray structure can inform on the protonation state of its amino acids is limited. The position of Asp213 is highlighted in Figure 1, as well as the amino acids in its neighborhood (Figure 1B). Asp213 is conserved in known mammalian Proteinase 3 sequences suggesting that it might have a functional role. It is positioned at the bottom of the S1 pocket and has been suggested to give it a polar character. Moreover, Asp213 lies in the vicinity of Asp194, which is a highly conserved residue in all mammalian chymotrypsin-like serine proteases. Asp194 makes a crucial hydrogen bond with the N-terminal amino acid (Ile16) in the mature form of these enzymes. The transition from the inactive zymogens to active enzymes requires the cleavage of the N-terminal dipeptide (Ala-Glu for PR3) by another enzyme. It leads to a structural rearrangement of the N-terminal region, which from an extended solvent-exposed conformation becomes inserted into the protein core and interacts with Asp194. This structural rearrangement leads to the catalytically active enzyme. This mechanism is common to all chymotrypsin-like serine proteases. Deprotonation of the N-terminal moiety and disruption of the Asp194-Nter salt-bridge at high pH in the homologous chymotrypsin lead to a loss of the catalytic activity.[34] The proximity of Asp 213 to the essential Asp194 and its conservation in the PR3 family raises the question of its functional role. Unfortunately, no experimental data have been reported on either the importance of Asp213 (e.g., from mutagenesis studies) or on its protonation state. Attributing the correct protonation state of this amino acid is crucial in our understanding of the stability and activity of Proteinase 3, and furthermore absolutely necessary for in silico ligand- and drug design strategies targeting PR3.

In what follows, we report the results of FDPB pK_a predictions of the aspartic acid 213 of Proteinase 3 using different sets of atomic coordinates, the X-ray structure and structures extracted from MD trajectories. MD simulations were performed using both forms of the aspartic acid (protonated and ionized) and in both the free enzymes and the ligand bound form. This ensures that several representative structures are sampled and minimizes bias due to simulation setup. pK_a predictions were then performed using varying values of protein dielectric constant ε_p. The influence of conformational sampling was evaluated by comparing the prediction obtained from average structures to those obtained from sampling the entire trajectories. Moreover, we describe the influence of the choice of the protonation state of Asp213 on substrate binding and active site structure of PR3, highlighting the importance of the choice of the correct protonation state to obtain reliable results from MD simulations. Finally results obtained with an empirical approach (propKa[22,23]) on the same coordinate sets are reported and compared to the results from the Poisson−Boltzmann approach.

This manuscript summarizes the difficulties met in predicting protonation states of buried functional amino acids and the questions that need to be addressed in the development of reliable titration protocols.

Methods

All calculations were performed using as starting point the X-ray structure of PR3 referenced 1FUJ (resolution 2.2 Å)[33] in the Protein Data Bank.[35,36] PR3 is crystallized as a tetramer with four structurally similar monomers. Experimental evidence points to the monomer and not the tetramer as the biologically active form of PR3. We therefore used only chain A in our calculations.[31] The monomer contains 221 amino acid residues. There are four disulfide bridges between cysteines 42−58, 136−201, 168−182, and 191−220. Histidine residues numbered 24, 48, 132 were protonated at their N^{ε} nitrogen atom, and the others (20, 40, 57, 71, 82, 147 and 179) at their N^{δ} nitrogen atom.

1. FDPB pK_a Calculations. The pK_a values were calculated using the UHBD program[37] and the PKALK package.[19]

Setup. We used a strategy similar to the one used by Schaefer et al.[19] The titratable residues considered were Asp, Glu, Arg, Lys, His, and Tyr. The titration calculations are based on the estimate of the electrostatic free energy for assembling the charge distribution of the protein with the residues either protonated or unprotonated ("two-states model"). The charge distribution and atomic radius of the titratable residues in their protonated and unprotonated forms were taken from the all-hydrogen parameter set of Charmm27.[39]

For Asp213 we considered two different models for the neutral form of Asp. In one model (the most common for acid titration) the proton is shared between the two oxygens. The charge of these oxygens are therefore reduced but no explicit positive charge is introduced.[38] Additionally, to check that the precise proton location did not influence the results, we tested a charge distribution for Asp213 where a proton is explicitly introduced on one of the oxygens.

Electrostatic Energy Calculations. The electrostatic interaction energies between the titration sites were calculated using the University of Houston Brownian Dynamics package (UHBD).[37] The program employs a solution of the linearized form of the Poisson−Boltzmann equation by the method of finite differences.[40] We then evaluate the change of the pK_a value of each titratable amino acid side chain when this side chain is moved from a hypothetical fully solvated state to its position in the protein. The final output is the calculated state of protonation at various pH values and pK_a shifts for all titratable sites in the protein. The protein dielectric boundary was taken to be a Richard's probe-accessible surface[41] computed with a spherical probe of radius 1.4 Å. We used different radius values of the probe but setting it to 2 or even 3 Å did not change the results significantly. A cubic $60 \times 60 \times 60$ grid point was used, the final grid spacing was 0.35 Å, and the number of surface dots per atom for calculating the solvent accessible surface was set to 1500. The ionic strength of the aqueous solution was set to 50 mM, and ion exclusion radius to 2.0 Å. The solvent dielectric constant was set to 80. Finally the temperature was set to 293 K and two distinct values for the solute dielectric constant were tested, 10 and 20. Water molecules were not included explicitly in the calculations. We did run calculations with a selection of explicit water molecules as well (results not shown, pK_a prediction on average structures) but it did not change significantly the predicted pK_a values.

Determination of Protonation State. There is an exponential dependence of the number of protonation states on the number of titrating sites. This makes statistically impossible the problem of finding the protonation state of our system that contains 46 titrable sites (42 on Proteinase3, 4 in the bound peptide). We thus used a Monte Carlo titration program[42] to determine the titration curve. The titration calculations for each site were performed as described in ref 19 and using shell scripts developed by M. Schaefer (personal communication). From the titration curve, the pK_a of each site can be derived as the pH where the proportion of protonated and unprotonated forms are equals. Following this strategy, it takes a total time of 1 CPU-hour on a standard Linux machine to calculate the pK_a values of the 46 titratable residues of a complex made of PR3 and a heptapeptide (sequence: Val-Ala-Asp-Val-Lys-Asp-Arg).

2. propKa Calculations. As a complement to the continuum electrostatic calculations described above, we also determined pK_a values with the propKa webserver.[22,23] The propKa method is an empirical procedure, where pK_a shifts from solution values are determined using empirically calibrated increments. The pK_a shift increment reflects the desolvation effects and intraprotein interactions that cause pK_a shift. A notable difference with continuum electrostatic method is that these increments are calibrated to reflect mostly the local environment of the titratable group. However, this method has proven to make extremely robust predictions in a large number of test cases.[22,23]

3. Molecular Dynamics Simulations. ***Setup.*** Two forms of Proteinase 3 were first simulated, one with a protonated carboxyl group on Asp213 (PR3-D^0) and the other one with the unprotonated carboxylate group (PR3-D^{-1}). The coordinates of the missing hydrogen atoms were determined using the HBUILD[43] facility in Charmm at pH 7. As a result, the total charge of PR3 was equal to +2 for PR3-D^0 and +1 for PR3-D^{-1}. To avoid simulations of non-neutral systems, ions were added to the enzyme-peptide complexes that had a net charge different from zero. The ions were added in the solvation boxes by replacing randomly chosen water molecules at a significant distance of the enzyme (>8 Å); no ions came in contact with the enzymes or peptides during the course of the simulations. Both PR3-D^0 and PR3-D^{-1} were also simulated with a peptide (sequence: Val-Ala-Asp-Val-Lys-Asp-Arg) designed to have a high affinity and specificity for PR3[31] and referred to as Popt in this manuscript. The peptide was docked into both forms of the enzyme as described previously.[31] We have run test calculations using either oxygen atoms of Asp213 (i.e., Oδ1 or Oδ2) as protonation site but it did not affect the results significantly; the orientation of the carboxylate and thus the average position of the protonated oxygen along the MD are identical in both cases.

MD Simulations Protocol. All MD simulations were run using NAMD[44,45] as described in ref 31. Beforehand, the systems were refined by geometry optimization, using the Charmm software (v30b1),[46] together with the Charmm27 force field[39] and then solvated in cubic boxes (60 Å) of pre-equilibrated water molecules using VMD.[47] After energy minimization, the solvated systems were progressively heated up to 300 K and equilibrated using periodic reassignment of the velocities to control the temperature. The systems were examined in the (N,P,T) ensemble, with Langevin dynamics to control the pressure and the temperature, SHAKE[48,49] to constrain the bond lengths between hydrogen and oxygen atoms of the water molecules and particle Mesh Ewald algorithm[50,51] for the long-range electrostatic forces. Once the systems had reached 300 K, they were stabilized during 100 ps. After the equilibration phase, production runs (i.e., without periodic reassignment of the velocities) were performed for 2000 ps. On the basis of the evolution of the RMSd between stored frames and the X-ray structure, we chose the sampling window to be between 500 and 2000 ps for all simulations (in total 1.5 ns).

Hydrophobic Contacts and Hydrogen Bonds. The presence of hydrogen bonds and hydrophobic contacts between the substrate and the enzyme is detected in the sampling windows of the simulations using Charmm, as described in detail in ref 31. In short, we consider a hydrogen bond to exist between an acceptor (A) and a donor (D) if the distance between A and D is lower than or equal to 2.4 Å,[52] and if the value of the angle formed by D−H−A is greater than or equal to 130°. A hydrophobic contact is assumed to exist when two candidate atoms are closer to each other than 3 Å. The occupancy of a hydrogen bond (or hydrophobic contact) is the ratio between its lifetime (sum of the duration of all events) and the sampling time (a hydrogen bond with an occupancy of 0.6 is present 60% of the simulation time).

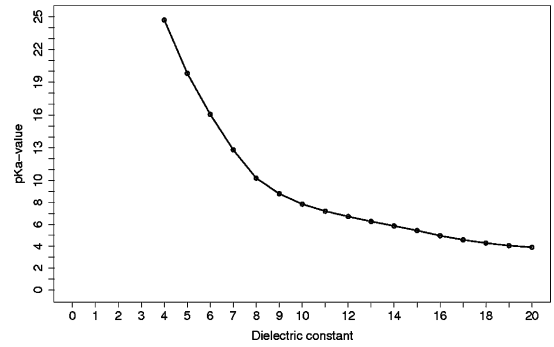

Figure 2. Variation of the predicted pK_a values with ε_p. The pK_a values (*Y*-axis) of Asp213 calculated from the atomic coordinates of the X-ray structure 1FUJ[33] (chain A) are reported for a range of dielectric constant from 4 to 21 (*X*-axis).

Results

1. pK_a Prediction from X-ray Structure. We first calculated the pK_a of Asp213 from the X-ray structure of PR3 (PDB ID 1FUJ[33]). Given the known dependency of the pK_a prediction on the choice of the protein dielectric constants (ε_p)[9,53,54] pK_a values of the titrable residues of PR3 were estimated with increasing ε_p values ranging from 1 to 20. While the predicted pK_a values for most of the titrable residues remain unchanged with the different ε_p values (data not shown), the predicted pK_a of Asp213 varies from values above 14 for low dielectric constants ($\varepsilon_p < 7$) down to values of about 3 for dielectric constants above 20 (see Figure 2). Thus, predicted pK_a values with ε_p comprised between 1 and 11 suggest that Asp213 is protonated, while the use of ε_p above 12 predicts a negatively charged carboxylate group. We compared two charge distributions for the neutral form of Asp213 (see Methods) but the results were not significantly different (Figure A, Supporting Information). The X-ray structure of PR3 shows a short distance between Oδ1 of Asp213 and the carbonyl oxygen atom of Gly197 and strongly suggests the presence of a hydrogen bond between the two atoms,[33] implying a protonated carboxyl group of Asp213. Although one is tempted to choose a low ε_p because this amino acid is rather buried, several studies have demonstrated increased performance of PB methods when using high dielectric constant for the protein.[8,15,19]

Buried water molecules can contribute to the stability of protein structures by forming hydrogen bonds with protein main-chain groups or stabilize unusual protonation states (e.g., ref 55). The explicit inclusion of buried water molecules in PB models has been shown to improve pK_a predictions.[56,57] However in our calculations, we could not see any significant changes of the predicted pK_a values between calculations with or without the crystal waters located in the vicinity of Asp213 (results not shown).

2. Active Site Conformations and pK_a Values along MD Simulations of PR3-D^0 and PR3-D^{-1}. We then studied the behavior of Asp213 and its environment along MD simulations of PR3. Simulations were launched that differed solely in the protonation state chosen for Asp213. The goal was 2-fold, (i) obtain conformational sampling for the pK_a prediction and (ii) evaluate the consequences of choosing one protonation state or another on the protein structure and dynamics. The simulations

TABLE 1: Hydrogen Bond Network in the Catalytic Triad for PR3-D0 and PR3-D- in Its Free Form and Bound to a Peptide[a]

		PR3-D^0			PR3-D^{-1}		
atom names	data	distance	angle	occupancy	distance	angle	occupancy
$N^{\varepsilon 2}{}_{His57}$-$O^{\gamma}{}_{Ser195}$	X-ray	2.36			2.36		
	MD	3.82 ± 0.43	112 ± 23	0.10	3.67 ± 0.36	122 ± 15	0.13
	MD Popt	2.87 ± 0.14	159 ± 11	0.97	3.66 ± 0.31	77 ± 23	0.05
$O^{\delta 1}{}_{ASP102}$-$N^{\delta 1}{}_{His57}$	X-ray	3.50			3.50		
	MD	2.85 ± 0.16	156 ± 13	0.94	3.21 ± 0.24	142 ± 12	0.50
	MD Popt	3.05 ± 0.35	151 ± 13	0.71	4.13 ± 0.25	126 ± 15	0.06
$O^{\delta 2}{}_{Asp102}$-$N^{\delta 1}{}_{His57}$	X-ray	2.54			2.54		
	MD	3.46 ± 0.36	138 ± 13	0.22	2.80 ± 0.12	151 ± 15	0.91
	MD Popt	3.21 ± 0.30	145 ± 11	0.58	3.46 ± 0.62	146 ± 15	0.51

[a] Average values (calculated along the sampling window) of distances (Å) and angles (°) are given together with their standard deviations. Distances between corresponding heavy atoms in the crystal are given for reference.

TABLE 2: Hydrogen Bond Network around Asp213 in Two Different Protonation States[a]

amino acid	atoms	X-ray	MD PR3-D^0	MD PR3-D^-
		distance (Å)	occupancy	occupancy
Asp194	$O^{\delta 2}$ ($H^{\delta 2}$)−O	3.2	0.28	-
Gly196	$O^{\delta 1}$−N (HN)	5.5	0.01	0.75
	$O^{\delta 2}$−N (HN)	7.1	-	0.29
Gly197	$O^{\delta 1}$−N (HN)	3.0	0.25	0.95
Ser195	$O^{\delta 2}$−O^{γ} (H^{γ})	6.9	-	0.94
WAT 5	$O^{\delta 1}$−O (H1, H2)	2.6	0.19, 0.20	0.25, 0.31
	$O^{\delta 2}$−O (H1, H2)	3.2	0.39 ($H^{\delta 2}$−O)	0.04, 0.01

[a] Occupancy of hydrogen bonds involving the carboxyl group ($O^{\delta 1}$ and $O^{\delta 2}$) of Asp213 in the simulations of PR3 complexed with a peptide ("-" = nondetected). Distances between corresponding heavy atoms in the crystal are given for reference.

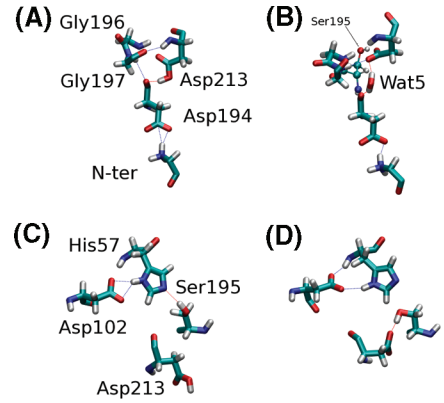

Figure 3. Hydrogen-bond networks observed in the simulations of PR3-D^0 and PR3-D^{-1} around Asp213 (A and B, respectively), and in the catalytic triad (C and D, respectively). The structures drawn are snapshots from the MD simulations with Popt bound to Proteinase 3. All amino acids are represented using licorice except Ser195 on picture B for which we used balls and sticks. The atoms are colored following their type (N, dark blue; C, cyan; O, red; H, white) and the dashed lines represent hydrogen bonds (cf. Tables 1 and 2).

were performed on the free protein (i.e., comparable to the X-ray structure without any ligand bound) and on PR3 complexed with Popt (see Methods section). The geometry of the enzyme active site and the environment of Asp213 were monitored during the simulations. Furthermore, we calculated the pK_a of Asp213 and all the other titratable amino acids for evenly distributed conformations along the trajectories (every 20 ps) as well as for the average structures calculated over the sampling windows. To account for the observed dependence of Asp213 on ε_p, both the values of 10 and 20 were used in the pK_a calculations.

Structure of the Catalytic Triad and the Environment of Asp213. The catalytic mechanism of PR3 is dependent on a specific and well-described network of hydrogen bonds in the active site of the enzyme.[58] We thus monitored these hydrogen bonds during the MD simulations. We report in Table 1 the average distances between amino acids of the catalytic triad along the four simulations. Furthermore, a list of hydrogen bonds between Asp213 and its partners in the two simulations with Popt are summarized in Table 2. The simulations of PR3-D^0 (i.e., with a neutral Asp213) bound to the peptide Popt reproduce the expected conformation of the catalytic triad; there are strong hydrogen bonds between histidine 57 and serine 195 and between histidine 57 and the aspartic acid 102. Oxygen atoms $O^{\delta 1}$ and $O^{\delta 2}$ of Asp102 exchange twice during the simulation, which explains that their average distance to the histidine N^{δ} is slightly different from the X-ray values. In the simulation of PR3-D^0 without a ligand bound, these hydrogen bonds are also strong (occupancy of 0.94 and 0.22 for $O^{\delta 1}$ and $O^{\delta 2}$, respectively). However the histidine-serine interaction is less prominent; the hydrogen bond is present only 10% of the sampling time compared to 97% when the ligand is present. In the latter case, because of steric hindrance the hydroxyl group is not free to rotate around the C^{α}−C^{β} bond and the lifetime of the hydrogen bond increases accordingly. Fujinaga and co-workers observed a short distance (2.98 Å) between a carboxylate atom of Asp 213 ($O^{\delta 2}$ in our atom name convention) and the carbonyl oxygen of Gly197 and suggest the formation of a hydrogen bond between these two groups (Asp213 would then be protonated). Although our simulation with protonated Asp213 reproduces this distance (PR3-D^0: $O^{\delta 1}$ 3.24 ± 0.51 Å, $O^{\delta 2}$ 4.5 ± 0.64 Å), we do not observe the formation of a strong hydrogen bond between the two moieties. Instead, the carbonyl group of Gly197 makes a strong hydrogen bond with the NH group of Asp213 (see Figure 3), $O^{\delta 2}$ (Asp213) interacts with the carbonyl oxygen of Asp194 or a water molecule, and the $O^{\delta 1}$ atom of Asp213 interacts relatively strongly with the NH group of Gly197. $O^{\delta 1}$ (nonprotonated oxygen, Asp213) interacts alternatively with NH (Gly197), the water molecule Wat5 (X-ray water), and occasionally its own backbone NH group (occupancy <5%).

We do not observe any significant difference in the hydrogen bond network within the catalytic triad between the two forms of PR3 (D^0 and D^{-1}) in the absence of ligand. However, the simulation of PR3-D^{-1} (i.e., with ionized Asp213) with the peptide bound reveals a highly perturbed catalytic triad. The hydroxyl oxygen O^{γ} of Ser195 interacts through a hydrogen bond with the carboxylate group of Asp213 instead of the nitrogen atom N^{ε} of His57 (occupancy 0.95, cf. Table 2); the distance between O^{γ}(Ser195) and $O^{\delta 2}$(Asp213) is thus decreased from a value of 7 Å in the X-ray structure to 2.8 Å. The hydrogen bonds between His57 and Asp102 weaken considerably (cf. Table 1). Moreover the geometry of the amino

TABLE 3: Influence of the Protonation State of Asp213 on the Orientation of Its Side Chain[a]

angle	X-ray	PR3-D^0	PR3-D^0	PR3-D^{-1}	PR3-D^{-1}
		free	Popt bound	free	Popt bound
$\chi1$	-52.4^a	-58.4 ± 8.1	-33.4 ± 54.5	-54.4 ± 7.2	51.8 ± 8.1
$\chi2$	47.7^b	24.3 ± 37.3	56.0 ± 63.2	56.2 ± 10.4	-78.6 ± 12.2

[a] $\chi1$ for monomer A in 1FUJ;[33] other monomers: $\chi1$(B) = $-56.3°$, $\chi1$(C) = $-31.7°$ and $\chi1$(D) = $-55.1°$. [b] $\chi1$ for monomer A in 1FUJ;[33] other monomers: $\chi2$(B) = 39.2°, $\chi2$(C) = 30.0° and $\chi2$(D) = 61.1°. [a] Comparison of the average values of the $\chi1$ and $\chi2$ angles during the MD simulations to the X-ray values (in degrees). Standard deviations are also reported.

acids around Asp213 is considerably perturbed compared to the X-ray structure. In particular we observe the formation of hydrogen bonds between the NH group of Gly196 and the carboxylate atoms O$^{\delta1}$ and O$^{\delta2}$ of Asp213, although the corresponding distances in the X-ray structure are 5.5 and 7.1 Å, respectively. Finally, the hydrogen bond between the carbonyl group of Gly197 and the NH group of Asp213 observed in the simulation with the neutral Asp 213 is virtually inexistent when Asp213 is not protonated.

As expected the protonation state of Asp213 affects also its interactions with buried water molecules. The water molecule Wat5 is hydrogen bonded to the carboxylate of Asp213; the distances O$^{\delta1}$−O and O$^{\delta2}$−O are 2.6 and 3.2 Å, respectively, in the X-ray structure. The simulation of PR3-D^0 with Popt lead to average distances of 3.3 ± 0.8 and 3.7 ± 0.8 Å, where we observe that the water molecule is either hydrogen bonded to O$^{\delta1}$ (with either of its hydrogen atoms H1 or H2) or to H$^{\delta2}$ through its oxygen atom. In the simulation with a negatively charged Asp213, these distances are much longer (4.0 ± 1.5 and 5.4 ± 1.3 Å) than the X-ray values and the total occupancy of hydrogen bonds between Wat5 and Asp213 is much lower than with a protonated Asp213.

These significant differences between the crystal structure and the simulations with the deprotonated Asp are also reflected in the orientation of the side chain of the aspartic acid. The differences in the values of the $\chi1$ (N−C$^\alpha$−C$^\beta$−C$^\gamma$) and $\chi2$ (C$^\alpha$−C$^\beta$−C$^\gamma$−O$^{\delta1}$) angles for the protonated and unprotonated forms of Asp213 illustrate the influence of the protonation state on the orientation of its side chain. Table 3 lists the X-ray values for $\chi1$ and $\chi2$ as well as the average values and standard deviations obtained from MD simulations (variation plots are given in Supporting Information, Figure B). The difference between protonated and unprotonated states is observed mostly when the peptide is bound to the enzyme. While in PR3-D^0, the average values of $\chi1$ and $\chi2$ are in the range of the ones found in the X-ray structure, there is a large deviation of the $\chi1$ and $\chi2$ angles in the simulation of PR3-D$^-$.

To summarize, when Asp213 is ionized and in the presence of a peptide bound in the enzyme, its side chain moves away from Gly197 and toward Ser195. The resulting interaction between Asp213 and Ser195 is strong with the formation of a hydrogen bond that persists for over 90% of the simulation time. This interaction competes with the crucial Ser-His hydrogen of the catalytic triad and would thus alter the proteolytic function of the enzyme. Interestingly this distortion is not observed in the absence of ligand, indicating that running MD simulations starting from the X-ray structure does not reveal the structural consequences of choosing an ionized Asp213. Moreover, as described below, it disturbs the peptide−enzyme interactions.

Interactions between PR3 and the Peptide Popt. In previous studies,[31,59] we described the interactions between PR3 and several peptidic substrates and identified the ligand-enzyme interaction pattern necessary to ensure efficient cleavage. Popt was designed to match this pattern and we demonstrated experimentally that it was efficiently and specifically cleaved by PR3. In a similar way, we inventoried the interactions (hydrogen bonds and hydrophobic contacts) observed along the MD simulations and estimated their lifetime for both forms of Asp213. Figure 4 displays the lifetime of the interactions between the residues of the ligand cleavable bond (P1-Val and P1′-Lys; labeled according to the Schechter and Berger convention[27]) and PR3.

When Asp213 is protonated (PR3-D^0, Figure 4A), the inventory of the interactions between PR3 and the residue P1

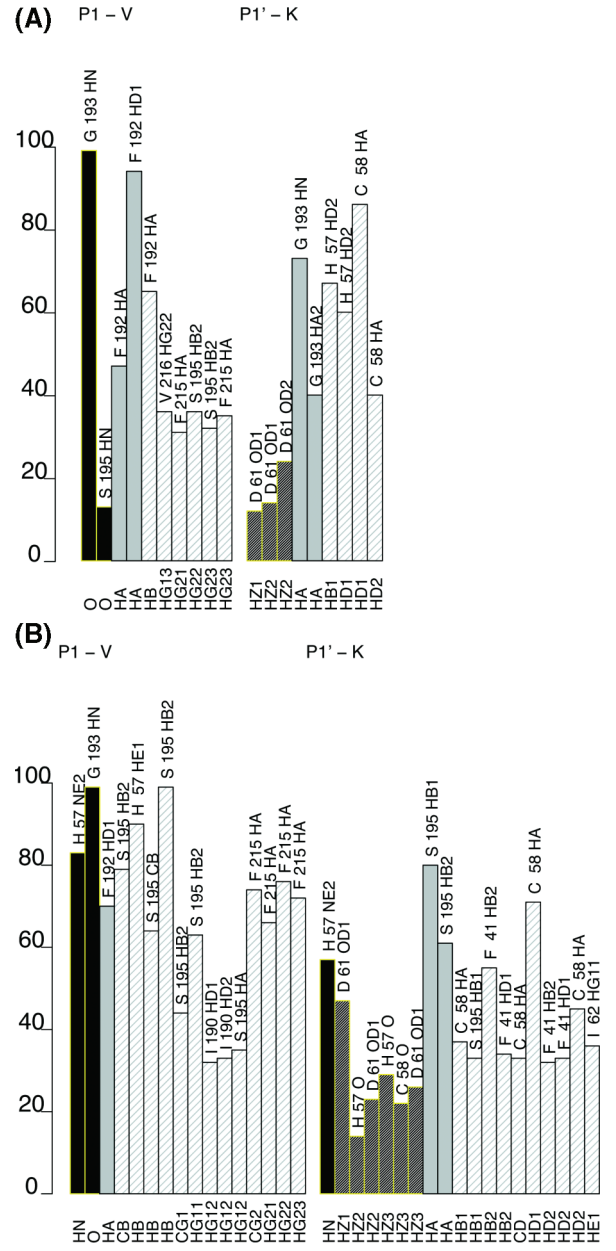

Figure 4. Interaction plot of PR3-D^0 and PR3-D^{-1} with the P1−P1′ aminoacids of Popt (Val-Lys). Hydrogen bonds and hydrophobic contacts between the Val-Lys cleavable segment of Popt and (A) PR3-D^0 and (B) PR3-D^{-1}; atoms involved (on *X*-axis) and interactions lifetime (on *Y*-axis). Black bars represent hydrogen bonds, gray bars represent hydrophobic contacts. Hashed bars signify interactions involving side-chains of the peptide, while nonhashed are used for interactions involving the backbone of the peptide (see refs 31 and 59).

of the peptide is in agreement with what is usually described for an effective positioning of the substrate with respect to the catalytic triad,[58,60] the presence of hydrophobic interactions between the P1 side chain and S1 (Val216, F192) and hydrogen bonds between the oxygen carbonyl of the P1 residue (oxyanion hole) and the backbone NH groups of Ser195 and Gly193. Asp213 in its neutral state is never seen to interact with the peptidic substrate Popt, which is in agreement with previously reported simulations where the peptide carried various P1 residues, among which even a polar P1 amino acid (Cys).[31] These results thus confirm the buried character of Asp213, which further indicate that a low value for the dielectric constant is appropriate for its titration.

In the simulations with the negatively charged form of Asp213 (PR3-D$^-$) the interaction between the carbonyl of the P1 residue (oxyanion hole) of the substrate and Ser195 (NH) of PR3 has vanished (Figure 4B) and the favorable interactions between the enzyme and the substrate are disrupted. Instead, a strong hydrogen bond is formed between the imidazole ring of His57 (N$^{\varepsilon 2}$) and the P1 and P1′ residues of the substrate (NH groups) lasting 83 and 55% of the simulation time, respectively. These hydrogen bonds compete with, or replace, the crucial ones between amino acids of the catalytic triad and in particular the Ser-His interaction (cf. Table 1).

pK_a Predictions of Asp213 along the Trajectory and on Average Structures. Every 20 ps along the 2 ns MD trajectories of both PR3-D^0 and PR3-D^{-1} free or in complex with Popt, the titration curve of Asp213 was calculated and a corresponding pK_a value derived. The values calculated for free PR3 with $\varepsilon_p = 10$, ligand-bound PR3 with $\varepsilon_p = 20$, and $\varepsilon_p = 10$ are plotted in Figure 5A−C, respectively.

The pK_a calculations yield very different results when the MD simulations are set up with Asp213 protonated (black curve) or ionized (red curve). In the latter case, the calculated pK_a is very low and fluctuates around negative values both with $\varepsilon_p = 20$ (Figure 5C) and 10 (Figure 5B) in the presence of Popt. For the free enzyme (Figure 5A), the pK_a predictions of Asp213 in PR3-D^{-1} are significantly higher than in the presence of peptide (Figure 5C). When Asp213 is in a protonated state and $\varepsilon_p = 20$ is used (Figure 5B, black curve), the resulting pK_a for its carboxylate group stays close to 4, which is approximately the value obtained for the free amino acid in solution. However, when $\varepsilon_p = 10$ is used (Figure 5A,C, black curves), the calculated pK_a is mostly above 7 and suggests a protonated, neutral state. The pK_a values were also estimated on average structures calculated along the sampling window and the values are reported in Table 4, and compared to the average values of the pK_a predictions calculated on each conformation in the corresponding trajectories (i.e., average of the values plotted on Figure 5B,C). We find the values to be comparable. The calculation of averaged titration curves in principle requires the use of a low dielectric constant.[14,15] The same calculations were performed with $\varepsilon_p = 4$ and led to extremely high values of the pK_a of Asp213 (>14) in simulations with a protonated Asp and extremely low values (<0) for PR3-D^{-1}. The pK_a obtained by averaging over the MD with a low dielectric constant (4) thus appear overestimated and the most realistic value correspond to averaging with $\varepsilon_p = 10$. These data confirm that when MD is run with the protonated Asp213, which maintains the X-ray structure of the active site, the Asp is predicted protonated. On the other hand, when the deprotonated state is used in the simulation, the structure does not conform to the X-ray data but the reorganization of the active site is such that the deprotonated state is favored in the

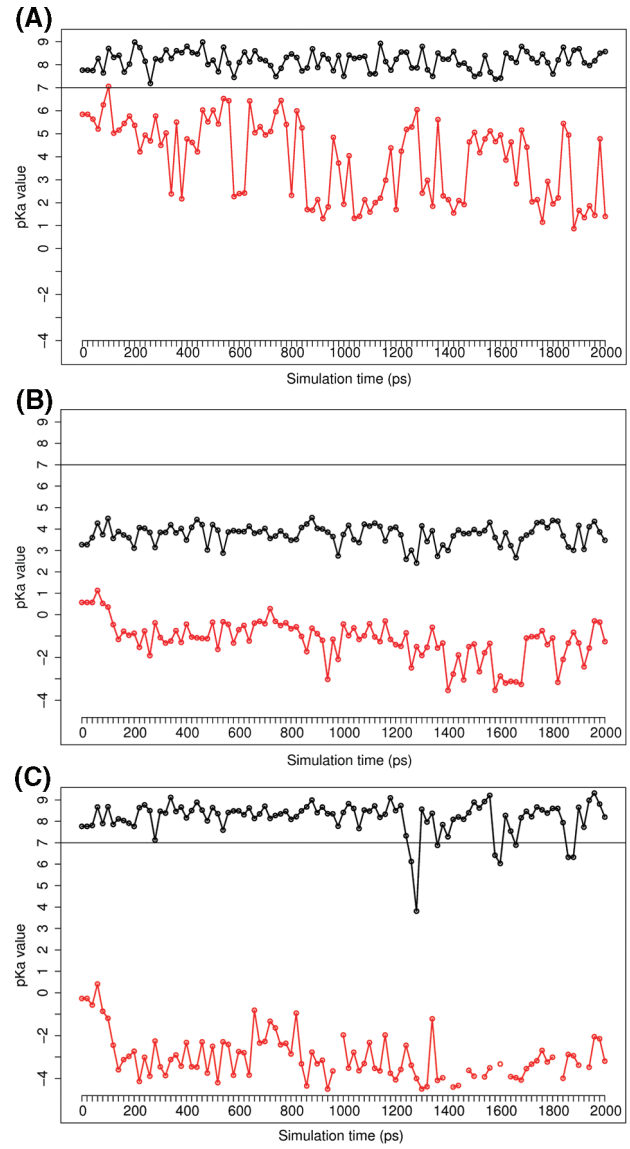

Figure 5. Asp213 pK_a values along the MD simulations of PR3-D^0 (black line) and PR3-D$^-$ (red line). Simulations of (A) free PR3, pK_a prediction with $\varepsilon = 10$, (B) ligand-bound PR3, pK_a prediction with $\varepsilon = 20$, and (C) ligand-bound PR3, pK_a prediction with $\varepsilon = 10$. pK_a calculations have been performed on conformations stored each 20 ps along the whole production run (0−2 ns). The horizontal black line correspond to the value where pK_a = pH = 7.

calculations. These results show that the titration curves evaluated along the MD trajectory depend not only on the value chosen for the dielectric constant but also on the initial protonation state chosen in the setup of the MD simulations. With $\varepsilon_p = 10$, an aspartic acid setup to its ionized form will be predicted to be ionized while it will be predicted to be protonated if it has been setup in this way. In other words, the pK_a prediction will lead to results matching the initial choice, as reported previously.[18,61,62]

Electrostatic Coupling between Asp213 and Other Charged Residues in PR3. We have used the average structures during the two simulations of PR3 complexed with Popt (PR3-D^0 and PR3-D^{-1}) to calculate the pK_a values of all the titratable residues of the enzyme (42 in total) for two values of the dielectric constant ($\varepsilon_p = 10$ and $\varepsilon_p = 20$). Whichever ε_p or Asp213 protonation state is used, none of the amino acids, except Asp213, yield a pK_a shifted enough to suggest a protonation state different from the one of the free amino acid in solution.

TABLE 4: pKa Values for Asp213 Estimated from Average Structures and from Trajectory Conformations[a]

A				
	$\varepsilon = 10$		$\varepsilon = 20$	
Popt bound	PR3-D^0	PR3-D^{-1}	PR3-D^0	PR3-D^{-1}
average structure	8.6	<0	4.2	4.3
trajectory/averaged	8.2 ± 0.8	<0	3.7 ± 0.4	<0
B				
	$\varepsilon = 10$		$\varepsilon = 20$	
PR3 free	PR3-D^0	PR3-D^{-1}	PR3-D^0	PR3-D^{-1}
average structure	8.4	4.9	4.3	2.3
trajectory/averaged	8.2 ± 0.4	3.9 ± 1.7	4.1 ± 0.3	2.1± 0.4

[a] (A) Comparison of the pKa values predicted on the average structures calculated from the simulations of PR3-D^0 and PR3-D^- (first row) complexed with Popt, and the average values computed from pKa predictions done on each conformation stored along the sampling window (second row). The corresponding values for PR3 unbound are given in (B).

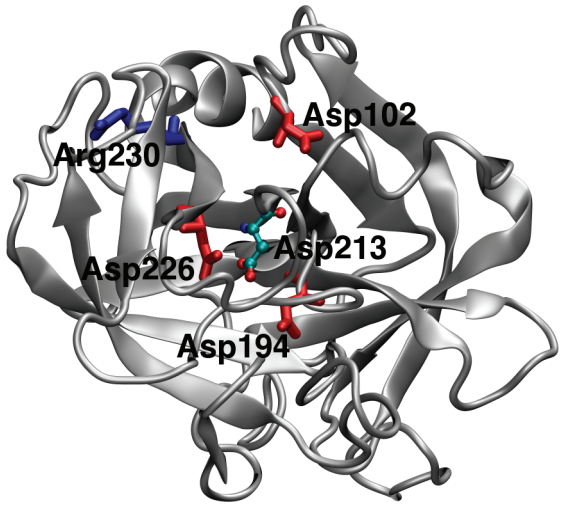

Figure 6. Titratable amino acids surrounding Asp213. The backbone of PR3 is displayed in gray using a cartoon representation to highlight its secondary structure elements. Asp213 and the titratable residues listed in Table 5 are displayed using a licorice representation and colored according to their type (acidic, red; basic, blue).

In fact, most of the 42 ionizable residues of PR3 have predicted pK_a's that are very similar to the one of the corresponding free amino acid in solution. The values for each amino acid are given in Supporting Information (Figure C).

To further decipher which particular interactions are mostly responsible for the shifted pK_a of Asp213, we artificially "discharged" the ionizable residues potentially having an electrostatic interaction with Asp213.[63] We considered amino acids within 9 Å of Asp213 in the average structure of PR3 from the MD trajectory of PR3-D^0:Popt (Asp102, Asp194, Asp 226, Arg230). Figure 6 shows their localization (X-ray structure). Their distance to Asp213 in the X-ray structure is reported in Table 5. Charged residues situated further than 9 Å did not affect the titration of Asp213 significantly (data not shown). For each of those residues, the partial charges were set to 0 and a new pK_a calculation was carried out on the average structure of the simulated complex between Popt and PR3-D^0. The difference between the resulting pK_a values, evaluated with $\varepsilon_p = 10$, and the value obtained with all charges "switched on" (8.6 and 8.4, with and without Popt, respectively, cf. Table 4) is reported in the third and fourth columns. The results are shown for $\varepsilon_p = 10$, but we observe the same trend with different values of ε_p (data not shown) with the exception of $\varepsilon_p = 20$ for which smaller shifts in the pK_a value of Asp213 are seen. The pK_a value of Asp213 is sensitive to the charges of the neighboring aspartic acids Asp102, Asp226, or Asp194 (Table 5). The largest modification is caused by the extinction of Asp194; it induces a downshift of 7 units of the pK_a of Asp213 in the case of PR3-D^0/Popt (3.8 in the case of PR3/D^0 without Popt) when its charges are turned off, which would lead to a different protonation state (i.e., charged) of the carboxylate group. Asp194 thus appears as a key residue in the environment of Asp213, as the two are strongly electrostatically coupled, and Asp194 has a key role in stabilizing the neutral Asp213 in both the presence and absence of ligand. The same strategy was applied to the average structures from the simulations of PR3-D^- free and bound to Popt (results not shown) and we observe the same trend (Asp194 has the biggest effect). The results show that Asp213 is predicted ionized, irrespective of the charge state of nearby residues. Extinguishing the charges of the peptide does not influence significantly the predicted pK_a values of Asp213, neither for PR3-D^0/Popt nor for PR3-D^{-1}/Popt.

TABLE 5: Influence of Other Polar Amino Acids on the pK_a Values of Asp213[c]

amino acid	distance to Asp213 (Å)	pK_a shift of Asp213[b] PR3-D^0:Popt	pK_a shift of Asp213[c]PR3-D^0
Asp102	9.0	−1	−0.6
Asp194	6.2	−7	−3.8
Asp226	7.8	−1.3	−1.2
Arg230	9.6	+0.2	+0.2

[a] List of amino acids of which the charges were extinguished, their distance to Asp213 (Ca−Ca distance) and the change of the predicted pK_a of Asp213 resulting from the extinction of their partial charges in simulations of PR3-D0 with and without the substrate. [b] Reference pK_a = 8.6. [c] Reference pK_a = 8.4.

TABLE 6: pKa Values for Asp213 Predicted by propKa[a]

X-ray		8.39
MD - Free enzyme	PR3-D^0	8.04
	PR3-D^{-1}	7.44
MD - PR3:Popt	PR3-D^0	8.46
	PR3-D^{-1}	3.87

[a] Comparison of the pKa values predicted on the X-ray structure and on the average structures from the simulations of free PR3-D^0 and PR3-D^- and complexed with Popt.

pK_a Prediction with propKa. We used propKa[22,23] to calculate the pK_a of Asp213 starting from the X-ray structure and the average structures calculated from the trajectories of PR3 with and without substrate and with both Asp213 ionized or protonated. The resulting values are reported in Table 6. Interestingly the pK_a value predicted from the X-ray structure (8.39) corresponds to a protonated Asp213, as expected from our MD simulations (and suggested by the X-ray structure). The pK_a shift is explained mostly by a large desolvation effect, an intraresidue backbone hydrogen bond (COO− with backbone NH) and a strong charge−charge interaction with Asp194. No hydrogen bond interaction with other amino acid side chains is identified by propKa. The same contributions are observed with the average structure of PR3-D^0/Popt. On the other hand, the pK_a value predicted from the MD average structure of PR3-D^-/Popt is 3.71; the most significant differences in the contributions used by propKa to this pK_a value and the pK_a shift calculated from the X-ray structure (or PR3-D^0/Popt) are the contribution from the Asp213-Ser195 hydrogen bond (cf. Table

2) and the loss of the Asp213-Asp194 charge−charge interaction. This contribution disappears because the distance between the C^γ of the carboxylate groups increases to 8 Å along the MD, compared to 5.5 Å in the X-ray structure. In the MD simulations of Proteinase 3 without a ligand bound and an ionized Asp213, we did not observe the Ser195-Asp213 hydrogen bond and the two aspartic acids (213 and 194) remain relatively close to each other ($C^\gamma-C^\gamma$ distance = 6 Å). The contribution to the pK_a shift calculated from propKa (8.04) thus reflects that lack of stabilizing hydrogen bond for an ionized carboxylate group, while still including the charge−charge interaction with Asp194.

Discussion and Conclusion

In this study, we investigated the protonation state of a buried aspartic acid (Asp213) in the human Proteinase 3. The amino acid is present in human PR3 (and its known mammalian homologues) but not in the human neutrophil elastase, and its position near the active site suggests that it could play a role in the function of these proteins, by a mechanism still to be discovered. As a first step in elucidating the role of Asp 213, we performed pK_a prediction using the PR3 X-ray structure as a starting point. Using continuum dielectric methods, which are commonly used for pK_a predictions, the results were highly dependent on the value chosen for the protein dielectric constant. Indeed calculations on the X-ray structure predict a protonated Asp213 for values of the protein dielectric constant below 11, and a charged state for higher ε_p values. This high sensitivity to the calculation setup was not observed for the other amino acid residues of PR3. This sensitivity of the results to the parameters of the calculations led us to make a detailed analysis of the parameters that influence Asp213 protonation and to devise a robust protocol for assessing its pK_a. We ran molecular dynamics simulation of the enzyme with both protonation states of the aspartic acid with and without a ligand bound in its active site. The structural analysis of the four simulations was confronted to the X-ray structure and to the data available on the reaction mechanism and ligand-binding pattern of these enzymes. In depth analysis of the simulations led us to conclude that a protonated state for Asp213 is the most coherent with the ensemble of available experimental data. Indeed, the structural requirements within the active site for an efficient cleavage of the peptidic substrate were not met in the simulations with an ionized carboxylate. Note that we needed simulations of the enzyme bound to a substrate to reach that conclusion; simulations of the X-ray structure (i.e., free enzyme) did not indicate strong structural perturbations of the active site when Asp213 was ionized. In the general case, this is however the only MD simulations that would be run (i.e., to gain conformational sampling starting from the X-ray coordinates set).[1] The publication describing the X-ray structure also hypothesized a protonated form of Asp213; however, the molecular dynamics indicate that the hydrogen bond network stabilizing the neutral acid differs from the network inferred for the crystal structure. Although in our case the structural and mechanistic information gave us robust indications to decide on the most appropriate protonation state, such detailed information is unfortunately seldom available. Instead, in many cases the prediction and description of the protein mechanism are the actual problem addressed by a study and require an ab initio prediction of the protonation state to be assigned to a given residue.

Conformational sampling, in particular through MD simulations, has been shown to improve pK_a predictions for a number of proteins.[14−16,62] We computed pK_a values of Asp 213 using MD trajectories. The pK_a values obtained from the titration curves calculated on the average structures are similar to values obtained on the corresponding averaged titration curves calculated along the trajectories, which is in agreement with the work of van Vlijmen et al.[15] Although one would tend to use a low protein dielectric constant for a buried amino acid residue, some studies have suggested that higher values led to optimum results.[8,15,19] In our study, the use of a protein dielectric constant below 10 is necessary to obtain pK_a predictions that are in agreement with the described reaction mechanism and ligand binding.

Provided that the optimum value of ε_p is used, the MD must be run with the right protonation state to lead to the prediction of the right protonation state. Others have discussed this bias earlier.[18,61,62] In the case of Proteinase 3, simulations run with a negatively charged Asp213 will lead to low pK_a values "confirming" the choice of the wrong protonation state in the MD setup. Calculations performed with propKa converged toward a protonated state of Asp213, except if we used as atomic coordinates the ones produced by MD runs setup with the wrong protonation state and the substrate bound. This emphasizes again the potential bias of conformational sampling, where the protein environment can relax to accommodate the wrong protonation state.

Using MD simulations, we have determined the protonation state of Asp213 of Proteinase 3, a necessary step for the development of specific ligands, inhibitors, and drugs targeting this therapeutic target in chronic inflammatory diseases. The pK_a value of Asp213, however, remains to be determined by experiments and we cannot conclude yet on the (potential) functional role of Asp213 from these calculations until mutagenesis experiments are performed. We however established that the use of a negatively charged Asp in computational studies of PR3 would lead to an unreliable structural model. Our results thus have an impact on future in silico studies aimed for example at finding low molecular weight compounds targeting PR3.

Furthermore we have highlighted the challenges in predicting the pK_a shift of a buried ionizable group in proteins using a finite difference Poisson−Boltzmann solver. More importantly this apparently trivial case illustrates the importance of carefully designing computational protocols for amino acid residues titration. An empirical method, significantly less demanding in terms of computer power leads to the same protonation state as predicted by our MD study; the contributions to the large pK_a shift of Asp213 in Proteinase 3 are a large desolvation effect, hydrogen bonds with the backbone of neighboring amino acids, an internal hydrogen bond and a large contribution from the proximity of another aspartic acid, Asp194. The latter contribution is identified both by the FDPB and the empirical approaches.

Acknowledgment. Funding for N.R. and E.H. was provided by the National Program for Research in Functional Genomics in Norway (FUGE) in the Research Council of Norway and the Bergen Science Foundation (Bergen Forskningsstiftelse). E.H. acknowledges an EU Grant MRTN-CT-2005-019335. Parallab (High Performance Computing Laboratory at the University of Bergen) and NOTUR (Norwegian consortium for high-performance computing) are thankfully acknowledged for provision of CPU time on their supercomputers. We thank Michael Schaeffer for fruitful discussions. A.D. thanks INSERM, CNRS, and Université de Strasbourg for financial support. N.R. and A.D. are thankful to the Egide AURORA program.

Supporting Information Available: Figure A: Dependence of the pK_a values of Asp213 on ε_p and the Asp model used. Figure B: Influence of the protonation state of Asp213 on the orientation of its side chain in MD simulations. Figure C: pK_a values for all titratable residues of PR3 and Popt (on MD average structures). This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Foloppe, N.; Sagemark, J.; Nordstrand, K.; Berndt, K. D.; Nilsson, L. *J. Mol. Biol.* **2001**, *310*, 449.
(2) Schiott, B.; Bruice, T. C. *J. Am. Chem. Soc.* **2002**, *124*, 14558.
(3) Celik, L.; Lund, J. D.; Schiott, B. *Biochemistry* **2007**, *46*, 1743.
(4) Cao, Z.; Mo, Y.; Thiel, W. *Angew. Chem., Int. Ed.* **2007**, *46*, 6811.
(5) Wlodek, S. T.; Antosiewicz, J.; McCammon, J. A. *Protein Sci.* **1997**, *6*, 373.
(6) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219.
(7) Yang, A. S.; Honig, B. *J. Mol. Biol.* **1993**, *231*, 459.
(8) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415.
(9) Demchuk, E.; Wade, R. C. *J. Phys. Chem.* **1996**, *100*, 17373.
(10) Alexov, E. G.; Gunner, M. R. *Biophys. J.* **1997**, *72*, 2075.
(11) Sham, Y. Y.; Chu, Z. T.; Warshel, A. *J. Phys. Chem. B* **1997**, *101*, 4458.
(12) Sham, Y. Y.; Muegge, I.; Warshel, A. *Biophys. J.* **1998**, *74*, 1744.
(13) Nielsen, J. E.; Vriend, G. *Proteins* **2001**, *43*, 403.
(14) Gorfe, A. A.; Ferrara, P.; Caflisch, A.; Marti, D. N.; Bosshard, H. R.; Jelesarov, I. *Proteins* **2002**, *46*, 41.
(15) van Vlijmen, H. W.; Schaefer, M.; Karplus, M. *Proteins* **1998**, *33*, 145.
(16) Zhou, H. X.; Vijayakumar, M. *J. Mol. Biol.* **1997**, *267*, 1002.
(17) Alexov, E. G.; Gunner, M. R. *Biochemistry* **1999**, *38*, 8253.
(18) Nielsen, J. E.; McCammon, J. A. *Protein Sci.* **2003**, *12*, 313.
(19) Schaefer, M.; Sommer, M.; Karplus, M. *J. Phys. Chem. B* **1997**, *101*, 1663.
(20) Lee, M. S.; Salsbury, F. R., Jr.; Brooks, C. L., III. *Proteins* **2004**, *56*, 738.
(21) Simonson, T.; Carlsson, J.; Case, D. A. *J. Am. Chem. Soc.* **2004**, *126*, 4167.
(22) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins* **2005**, *61*, 704.
(23) Bas, D. C.; Rogers, D. M.; Jensen, J. H. *Proteins* **2008**, *73*, 765.
(24) Ondrechen, M. J. *Current Protocols in Bioinformatics*; Wiley: New York, 2004; Chapter 8, Unit 86.
(25) Murga, L. F.; Wei, Y.; Ondrechen, M. J. *Genome Inform.* **2007**, *19*, 107.
(26) Wei, Y.; Ko, J.; Murga, L. F.; Ondrechen, M. J. *BMC Bioinf.* **2007**, *8*, 119.
(27) Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157.
(28) Korkmaz, B.; Attucci, S.; Hazouard, E.; Ferrandiere, M.; Jourdan, M. L.; Brillard-Bourdet, M.; Juliano, L.; Gauthier, F. *J. Biol. Chem.* **2002**, *277*, 39074.
(29) Koehl, C.; Knight, G.; Bieth, J. *J. Biol. Chem.* **2003**, *278*, 12609.
(30) Korkmaz, B.; Attucci, S.; Moreau, T.; Godat, E.; Juliano, L.; Gauthier, F. *Am. J. Respir. Cell Mol. Biol.* **2004**, *30*, 801.
(31) Hajjar, E.; Korkmaz, B.; Gauthier, F.; Brandsdal, B. O.; Witko-Sarsat, V.; Reuter, N. *J. Med. Chem.* **2006**, *49*, 1248.
(32) Korkmaz, B.; Hajjar, E.; Kalupov, T.; Reuter, N.; Brillard-Bourdet, M.; Moreau, T.; Juliano, L.; Gauthier, F. *J. Biol. Chem.* **2007**, *282*, 1989.
(33) Fujinaga, M.; Chernaia, M. M.; Halenbeck, R.; Koths, K.; James, M. N. *J. Mol. Biol.* **1996**, *261*, 267.
(34) Fersht, A. R. *J. Mol. Biol.* **1972**, *64*, 497.
(35) Abola, E.; Bernstein, F. C.; Bryant, S. H.; Koetzle, T. F.; Weng, J. Protein Data Bank. In *Crystallographic Databases-Information Content, Software Systems, Scientific Applications*; Data Commission of the International Union of Crystallography: Bonn, Germany, 1987.
(36) Bernstein, F. C.; Koetzle, T. F.; Willams, G. J. B.; Meyer, E. F. J.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535.
(37) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. *Comp. Phys. Comm.* **1991**, *62*, 187.
(38) Bashford, D. *Front. Biosci.* **2004**, *9*, 1082.
(39) Foloppe, N.; MacKerell, A. D. *J. Comput. Chem.* **2000**, *21*, 86.
(40) Garrett, A. J. M.; Poladian, L. *Ann. Phys.* **1988**, *188*, 386.
(41) Richards, F. M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151.
(42) Beroza, P.; Fredkin, D. R.; Okamura, M. Y.; Feher, G. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5804.
(43) Brunger, A. T.; Karplus, M. *Proteins* **1988**, *4*, 148.
(44) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comp. Phys.* **1999**, *151*, 283.
(45) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
(46) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(47) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33.
(48) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(49) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24.
(50) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(51) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(52) Deloof, H.; Nilsson, L.; Rigler, R. *J. Am. Chem. Soc.* **1992**, *114*, 4028.
(53) King, G.; Lee, F. S.; Warshel, A. *J. Chem. Phys.* **1991**, *95*, 4366.
(54) Wisz, M. S.; Hellinga, H. W. *Proteins* **2003**, *51*, 360.
(55) Likic, V. A.; Juranic, N.; Macura, S.; Prendergast, F. G. *Protein Sci.* **2000**, *9*, 497.
(56) Gibas, C. J.; Subramaniam, S. *Biophys. J.* **1996**, *71*, 138.
(57) Fitch, C. A.; Karp, D. A.; Lee, K. K.; Stites, W. E.; Lattman, E. E.; Garcia-Moreno, E. B. *Biophys. J.* **2002**, *82*, 3289.
(58) Hedstrom, L. *Chem. Rev.* **2002**, *102*, 4501.
(59) Hajjar, E.; Korkmaz, B.; Reuter, N. *FEBS Lett.* **2007**, *581*, 5685.
(60) Perona, J. J.; Craik, C. S. *Protein Sci.* **1995**, *4*, 337.
(61) Bashford, D.; Gerwert, K. *J. Mol. Biol.* **1992**, *224*, 473.
(62) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *Biochemistry* **1996**, *35*, 7819.
(63) Dejaegere, A.; Choulier, L.; Lafont, V.; De Genst, E.; Altschuh, D. *Biochemistry* **2005**, *44*, 14409.

JP902930U