PROTEINS: Structure, Function, and Genetics 47:469–480 (2002)

# Simulating Proteins at Constant pH: An Approach Combining Molecular Dynamics and Monte Carlo Simulation

**Roland Bürgi,[1] Peter A. Kollman,[2] and Wilfred F. van Gunsteren[1]***
[1]*Laboratory of Physical Chemistry, Swiss Federal Institute of Technology Zürich, Zürich, Switzerland*
[2]*Department of Pharmaceutical Chemistry, University of California, San Francisco, California*

***ABSTRACT*** **For the structure and function of proteins, the pH of the solution is one of the determining parameters. Current molecular dynamics (MD) simulations account for the solution pH only in a limited way by keeping each titratable site in a chosen protonation state. We present an algorithm that generates trajectories at a Boltzmann distributed ensemble of protonation states by a combination of MD and Monte Carlo (MC) simulation. The algorithm is useful for pH-dependent structural studies and to investigate in detail the titration behavior of proteins. The method is tested on the acidic residues of the protein hen egg white lysozyme. It is shown that small structural changes may have a big effect on the $pK_A$ values of titratable residues. Proteins 2002;47:469–480.**


**Key words: molecular dynamics; Monte Carlo; protein titration; constant pH; GROMOS; lysozyme**

## INTRODUCTION

It has been known for a long time that the structure and function of proteins are pH dependent.[1–4] The pH dependence of structure and function is induced by changes in the titration behavior of the ionizable (titratable) residues of the protein at different pH values. The factors that influence the titration behavior of an ionizable residue are the pH of the surrounding solution and the interactions of the residue in question with its local environment, that is, the protonation state of the other titratable residues and the structure of the protein. These contributions are tightly coupled: changing the titration behavior of one residue can change the titration behavior of the other ionizable residues and the structure of the protein in a hardly predictable way. This makes titration curves of solvated proteins difficult to interpret and the experimental investigation of the titration behavior of residues in a protein nontrivial.[5–7]

There are two aspects that make molecular dynamics (MD) simulations of proteins at constant pH desirable: the investigation of the structure and function of the protein as a function of pH, and the analysis of the titration behavior of its ionizable residues. Although in recent years a number of methods have been developed to calculate the $pK_A$ values of ionizable residues, only a few approaches to account for the pH in an MD simulation have been reported. Most of the current MD simulations account for the pH through the choice of a particular, fixed protonation state of the protein. However, this model has flaws, particularly in the case when ionizable residues are close in space and when the $pK_A$ of a site is close to the chosen pH. In such a model, the ensemble of protonation states is replaced by one state only, whereas it is well known from experiments that many ionizable groups in a protein have their $pK_A$ values significantly changed from the standard values for the corresponding single amino acid in solution, which implies a dependence of the protonation state upon protein structure. Mertz and Pettitt[8] introduced a grand canonical Lagrangian into MD simulation that allows for protons to be exchanged between molecules. The application of this algorithm to a protein in water would require a potential energy function that accurately represents neutral, protonated, and deprotonated water molecules. More recently, methods have been developed by Baptista et al.,[9] in which a potential of mean force is used to account for an implicit titration, and by Börjesson and Hünenberger,[10] in which the protonation state for each titration site can vary continuously. Both these methods, however, are based on a mean field approximation: the protonation state is treated as a parameter that describes a continuous change from protonation to deprotonation. The value of this parameter is determined by minimizing the electrostatic energy of an ionizable site. Thus, the various protonation states are not sampled, but the protonation state follows adiabatically the variation in the potential at the ionizable site. The description of a protonation state through a continuous parameter is not equivalent to treating the protonation states explicitly, because physical observables can be affected by the protonation state of a residue in a highly nonlinear way: as soon as there is a strong coupling between the titratable residues, the titration curves get significantly distorted,[11] and it seems necessary to include the neutral and charged form of residues explicitly rather

Grant sponsor: Schweizerischer Bundesaurt für Bildung und Wissenschaft; Grant number: C00.036 (COST 09/0014/98).

*Correspondence to: Wilfred F. van Gunsteren, Laboratory of Physical Chemistry, Swiss Federal Institute of Technology Zürich, CH-8092 Zürich, Switzerland. E-mail: wfvgn@igc.phys.chem.ethz.ch

Received 5 June 2001; Accepted 27 September 2001

than representing the protonation state through intermediate or extrapolated charge states.[12]

The current methods to analyze the titration behavior of ionisable sites depend on determining the free energy difference of protonating the site ($\Delta F$) and relating it to a $pK_A$ value through the relation

$$\Delta F = \log(10) \cdot k_B T(\mathrm{pH} - \mathrm{pK_A}) \quad (1)$$

where $k_B$ is the Boltzmann constant and $T$ is the temperature of the system.[13] Into the calculation of $\Delta F$ enter a classical ($\Delta F^{\mathrm{class}}$) and a quantum-mechanical ($\Delta F^{\mathrm{quant}}$) contribution. $\Delta F^{\mathrm{class}}$ is caused by the change of nonbonded (electrostatic and van der Waals) interactions and the change of bond lengths and bond angles at the ionizable site. $\Delta F^{\mathrm{quant}}$ consists of the free energy difference of forming or breaking the bond to the proton and the free energy difference of (de)solvating the proton in water. Because $\Delta F^{\mathrm{quant}}$ cannot be calculated classically, classical methods attempt to calculate $pK_A$ shifts rather than absolute $pK_A$ values by calculating $\Delta F^{\mathrm{class}}$ for a protein residue ($\Delta F^{\mathrm{class}}_{\mathrm{prot}}$) versus $\Delta F^{\mathrm{class}}$ for the corresponding amino acid solvated singly in water ($\Delta F^{\mathrm{class}}_{\mathrm{ref}}$),

$$\Delta F^{\mathrm{class}}_{\mathrm{shift}} = \Delta F^{\mathrm{class}}_{\mathrm{prot}} - \Delta F^{\mathrm{class}}_{\mathrm{ref}} \quad (2)$$

The quantum-mechanical contributions are assumed to be the same for the reference compound as for the residue in the protein, and thus the resulting free energy difference ($\Delta F^{\mathrm{class}}_{\mathrm{shift}}$) can be equated to the total free energy shift,

$$\Delta F_{\mathrm{shift}} = \Delta F^{\mathrm{class}}_{\mathrm{shift}} \quad (3)$$

$\Delta F_{\mathrm{shift}}$ can then be related to the $pK_A$ shift of a residue in a protein with respect to its corresponding single amino acid,

$$\mathrm{pK_A} = \mathrm{pK_{A,ref}} + \mathrm{pK_{A,shift}}. \quad (4)$$

It appears that the relevant part of $\Delta F_{\mathrm{shift}}$ is mainly caused by the change of the electrostatic interactions with the residue. The model by Linderstrøm-Lang[14] incorporates two contributions to the electrostatic potential: the interactions of the residue with dipolar groups and charged residues within the protein, and the interaction of the residue with the solvent. Most of the current methods to investigate electrostatic effects in proteins are based on solving numerically the linearized Poisson-Boltzmann (PB) equation,[15] thus modeling a continuum solvent around the protein. A variety of difficulties is encountered when calculating $pK_A$ values by solving the PB equation. On the one hand, there are several parameters for which a value has to be chosen or which can be determined by fitting to a given set of $pK_A$ values: the partial atomic charges and van der Waals radii of atoms[16,17] and values for the dielectric constant inside and outside the protein.[18–21] On the other hand, the accuracy of a continuum solvent model can be questioned, especially for protonation sites close to the continuum solvent. There, the choice of the computational grid comes to play: whether a grid point lies just within the protein or just outside the protein can affect the calculated $pK_A$ values significantly. Efforts have been made to alleviate this problem by introducing single explicit water molecules.[12,22] In addition, calculations based on a single structure of the protein neglect structural fluctuations and seem to be less accurate than calculations based on ensembles.[23] Alternative methods to compute $pK_A$ values are based on microscopic polarization models[24,25] or the analysis of Gaussian fluctuations.[26] A recent review comparing different methods for calculating electrostatic energies and $pK_A$ values of proteins can be found in Ref. 27.

The calculation of $\Delta F_{\mathrm{shift}}$ through Eq. 2 is very demanding for the accuracy of methods to calculate $\Delta F^{\mathrm{class}}_{\mathrm{prot}}$ and $\Delta F^{\mathrm{class}}_{\mathrm{ref}}$: the latter are both of the order of $10^2$ kJ/mol, whereas $\Delta F_{\mathrm{shift}}$, the relevant free energy difference, is of the order of $10^0$ kJ/mol. Thus, the errors of the calculations of $\Delta F^{\mathrm{class}}_{\mathrm{prot}}$ and $\Delta F^{\mathrm{class}}_{\mathrm{ref}}$ must be well below 1%. Therefore, the various approximations made in the methods described above—only considering electrostatic interactions, use of a mean-field solvent, lack of an appropriate ensemble of structures, and so forth—may reduce the accuracy of the calculated $pK_A$ values significantly. Even if the free energy differences are averaged over an ensemble of structures, this ensemble is usually generated by MD simulation of the protein at one protonation state only, which induces the problem of implying a chosen protonation state in the calculations. Such a procedure does not account for structural changes of the protein as a function of its protonation state.

In the present work, we propose an approach for constant pH simulations combining MD simulation of the canonical ensemble of protein structures at one protonation state with Monte Carlo (MC) simulation of the grand canonical ensemble of protonation states of the protein. The method allows for the generation of structures of the protein at a Boltzmann distributed ensemble of protonation states. Thus, it is possible to analyze the structures of the protein as a function of the pH. Furthermore, many of the problems of the existing methods to analyze the titration behavior of ionizable residues are reduced: the free energy differences of protonating and deprotonating the titration sites can be averaged over a correctly weighted ensemble of structures and protonation states, and the free energy differences are calculated by including all the changes in interaction energy due to variations in protein structure and solvent configurations.

## THEORY AND METHODS

### Outline of the Algorithm

The algorithm presented in this work is based on MC simulations in the grand canonical ensemble of protonation states. At each MC step, a change in protonation state of a titratable residue is attempted. The free energy difference ($\Delta F$) of protonating or deprotonating the residue is calculated by MD simulation in phase space through free energy perturbation (FEP) theory. For protonating a residue, $\Delta F$ is given by

$$\Delta F = \log(10) \cdot k_B T(\mathrm{pH} - \mathrm{pK_{A,ref}}) + \Delta F^{\mathrm{class}}_{\mathrm{prot}} - \Delta F^{\mathrm{class}}_{\mathrm{ref}} \quad (5)$$

where $k_B$ is the Boltzmann constant, $T$ is the temperature of the system, pH is a simulation parameter, $pK_{A,ref}$ is the

$pK_A$ of the corresponding reference compound, $\Delta F_{prot}^{class}$ is the free energy difference as calculated for the residue in the protein, and $\Delta F_{ref}^{class}$ is the free energy difference as calculated for the residue in the reference compound. The MC step is accepted according to the probability $p$ of the Metropolis criterion:

$$p = \begin{cases} 1 & \Delta F \le 0 \\ e^{-\Delta F/k_B T} & \Delta F > 0 \end{cases} \quad (6)$$

Note that, because the MC algorithm affects the grand canonical ensemble of protonation states, the potential used in the Metropolis criterion Eq. (6) is the free energy $F$ rather than the mean energy $E$. Into the calculation of $\Delta F$ enter not only contributions due to the change in protonation but also contributions due to conformational changes induced by the change in protonation state. These contributions are desired, because the purpose of the algorithm is to generate properly weighted conformations at a given pH.

The choice of the initial structures for the MC steps as well as the generation of the total simulation trajectory is described in detail in Generation of the Trajectories.

## Test Systems

The method was tested on hen egg white lysozyme (HEWL). To speed up the calculations and to improve the sampling, none of the simulations were performed at a high pH; the system was simulated at pH values of 4, 3, and 2. Thus, the MC simulations considered only protonation states of the negatively charged residues (Asp, Glu) and the carboxy terminus, and all the positively charged residues (Arg, His, Lys) and the N terminus were fully protonated.

As reference compounds for aspartic acid (Asp) and glutamic acid (Glu), single blocked amino acids (Me—CO—NH—X—CO—NH—Me) were chosen. For the C-terminal carboxyl group (Cter), a single alanine with a blocked N-terminal group (Me—CO—NH—Ala—COOH) served as a reference. The $pK_A$ values for the reference compounds were taken from Ref. 28: 4.0 for Asp, 4.4 for Glu, and 3.8 for Cter.

## Molecular Model and Simulation Setup

All MD simulations were carried out by using the simulation package GROMOS96.[29,30] All reference compounds (Asp and Glu peptides and the C-terminal peptide Ala) and the protein HEWL were modeled by using the GROMOS96 biomolecular force field, parameter set 43A1.[29] The parameters for the carboxylic acid groups were chosen as described in Ref. 31. Water was modeled by using the simple point charge (SPC) model.[32]

As the starting structure for HEWL, the Brookhaven Protein Data Bank (PDB) entry 1AKI obtained at pH = 4.48 was chosen. To obtain the starting structure for the reference compounds, an energy minimization in vacuo was carried out starting from a randomly scattered distribution of atoms. The reference compounds as well as HEWL were solvated in a periodic truncated octahedron. The number of water molecules solvating the reference compounds Asp, Glu, and Cter were 1267, 1311, and 1238, respectively. HEWL was solvated in 5956 water molecules. The volume of the truncated octahedron was kept constant in all simulations, the distance of the solute to the square planes of the truncated octahedra being 1.8 nm for the reference compounds and 1.2 nm for HEWL. The temperatures of solute and solvent were separately weakly coupled to a temperature bath of 298 K with a coupling time of 0.1 ps.[33] All bond lengths were constrained by using the SHAKE algorithm[34] with a relative geometry precision of $10^{-4}$.

The nonbonded interactions were calculated with twin-range cutoff radii of 0.8 and 1.4 nm. In the intermediate range, the nonbonded forces were calculated every five timesteps of 2 fs. For the long-range electrostatic interactions, a reaction field[35] was used with the dielectric constant $\varepsilon_{RF} = 54$.[36]

### *Counter ions*

In the algorithm presented in this work, the total charge is not kept constant. It would be desirable to find a way to resolve this problem. However, the addition of charged salt ions is not at all trivial: adding counter ions to an MD simulation is a widely discussed topic.[37,38] In the present formalism, counter ions would have to be introduced so that they are charged when the corresponding residue is charged and uncharged otherwise. Whenever a residue changes its protonation state from uncharged to charged, the corresponding counter ion would essentially be located at a random position.

In the present study we chose to neglect the change in total charge rather than having to deal with unpredictable effects due to badly located counter ions. Of course, this problem has to be treated, and a solution has to be found to keep the total charge constant.

## Free Energy Calculations

The dependence of the Hamiltonian $H(\lambda)$ on $\lambda$ is given in Refs. 29 and 30. The deprotonated and protonated Hamiltonians, that is, the force field parameters of both states, are defined in Ref. 31. For the nonbonded interactions of the perturbed atoms, a soft core interaction was used[29,30,39] with the soft core parameters $\alpha_{LJ} = 0.5$ and $\alpha_C = 0.5$ nm$^2$.

### *Free energy calculations for the reference compounds*

All the reference free energy differences were calculated with the thermodynamic integration (TI) method.[40,41] For 18 values of the coupling parameter $\lambda$ ($\lambda$ = 0, 0.03, 0.07, 0.085, 0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.7, 0.8, 0.85, 0.9, 0.93, 0.97, 1), the system was equilibrated for 2.5 ps, and the derivative $\partial H/\partial\lambda$ was sampled for a period of 97.5 ps. The value of the derivative $\partial H/\partial\lambda$ was saved every 0.5 ps. The measurements were sufficiently uncorrelated for the standard deviation of $\partial H/\partial\lambda$ to be a suitable measure of the error.

The integration of $\langle \partial H/\partial\lambda \rangle_\lambda$ over $\lambda$ was carried out by performing a polynomial fit, weighing the different points with their respective standard deviations. The error of the

**TABLE I. Free Energy Differences $\Delta F^{class}$ of Protonation in (kJ/mol) for the Three Reference Compounds (Asp, Glu, Cter) and for Three Sites in HEWL (7Glu, 18Asp, 35Glu) as Obtained by Using Different Protocols[†]**

| Method | Asp | Glu | Cter | 7Glu | 18Asp | 35Glu |
|---|---|---|---|---|---|---|
| TI | 206.9 | 212.8 | 239.4 | 211.0 | 210.2 | 174.1 |
| $SG_{20}$ | 221.7 | 212.9 | 240.9 | 244.2 | 230.0 | 181.3 |
| $\Delta SG_{20}$ | 7.2 | 0.0 | 0.6 | 15.7 | 9.4 | 4.2 |
| $SG_{20,rev}$ | 208.9 | 201.0 | 239.3 | 211.2 | 210.8 | 159.3 |
| $\Delta SG_{20,rev}$ | 1.0 | 5.6 | 0.0 | 0.1 | 0.3 | 8.5 |
| $SG_{20,av}$ | 215.3 | 207.0 | 240.1 | 227.7 | 220.4 | 170.3 |
| $\Delta SG_{20,av}$ | 4.1 | 2.7 | 0.2 | 7.9 | 4.8 | 2.2 |
| $SG_{40}$ | 208.0 | 225.6 | 240.8 | 213.8 | 231.2 | 190.2 |
| $\Delta SG_{40}$ | 0.6 | 6.0 | 0.6 | 1.3 | 10.0 | 9.2 |
| $TI_1$ | 220.3 | 225.1 | 255.8 | 215.7 | 227.0 | 194.3 |
| $\Delta TI_1$ | 6.5 | 5.7 | 6.9 | 2.2 | 8.0 | 11.6 |
| $TI_{1,rev}$ | 203.1 | 208.9 | 220.8 | 215.0 | 206.5 | 147.2 |
| $\Delta TI_{1,rev}$ | 1.8 | 1.8 | 7.8 | 1.9 | 6.3 | 15.4 |
| $TI_{1,av}$ | 211.7 | 217.0 | 238.3 | 215.3 | 216.7 | 170.8 |
| $\Delta TI_{1,av}$ | 2.3 | 2.0 | 0.5 | 2.0 | 3.1 | 1.9 |
| $TI_5$ | 210.5 | 216.0 | 240.7 | 196.1 | 212.9 | 178.9 |
| $\Delta TI_5$ | 1.8 | 1.5 | 0.5 | 7.1 | 1.3 | 2.8 |

[†]Thermodynamic integration (TI), slow growth during 20 ps forward ($SG_{20}$) and during 20 ps backward ($SG_{20,rev}$), the average of the forward and backward results ($SG_{20,av}$), slow growth during 40 ps ($SG_{40}$), TI with 1 ps windows in the forward ($TI_1$) and backward ($TI_{1,rev}$) direction, the average of the forward and backward results ($TI_{1,av}$), and TI with 5-ps windows ($TI_5$). The relative error of method $X$ with respect to the TI results, denoted as $\Delta X$ is given in percent.

integral was estimated by a Monte Carlo integration: The value of $\langle \partial H/\partial\lambda \rangle_\lambda$ was randomly chosen from the normal distribution (defined by the mean value and standard deviation of $\partial H/\partial\lambda$) at each $\lambda$ value, and the error of the integration was taken as the standard deviation of the mean value of 100 such integrals. For all the reference compounds, the error of the free energy differences as calculated by TI is about 1.5 kJ/mol, which is below 1%. The results are presented in the first row and the first three columns of Table I.

### *Free energy calculations for HEWL*

Because the accurate free energy calculation using TI requires 1.8 ns of simulation time per $\Delta F^{class}$ value, this method is computationally too costly for the calculation of the free energy differences to be used in the MC steps. A variety of fast one- and two-step (using one or two $\lambda$-values) approximate methods have been tested on aspartic acid.[31] The method with the best results presented there, linear response theory, yields results within an error range of around 10% for changes in one protonation state. However, because the relevant free energy differences $\Delta F$ are of the order of 1% of the calculated free energy differences $\Delta F^{class}_{prot}$ and $\Delta F^{class}_{ref}$, and $\Delta F$ occurs in the exponent of the Metropolis criterion Eq. (6), a more accurate (multistep) method is needed to calculate $\Delta F^{class}_{prot}$. Two multistep methods were tested for each of the reference compounds as well as for three sites in HEWL (7Glu, 18Asp, and 35Glu).

***Slow growth.*** In the slow growth (SG) procedure,[42–44] the value of $\lambda$ is changed every timestep by a small amount. Instead of an ensemble average, one measurement of $\partial H/\partial\lambda$ per $\lambda$-point is used for the integration. The value of $\partial H/\partial\lambda$ was saved every 0.1 ps. The integration was carried out by fitting an eighth order polynomial through the measured points and integrating the polynomial analytically. Two SG simulations were carried out over 20-ps periods for both the forward ($\lambda = 0 \rightarrow 1$) and backward ($\lambda = 1 \rightarrow 0$) change in $\lambda$ ($SG_{20}$ and $SG_{20,rev}$). The results are listed, together with their relative errors to the TI results, in the second to fifth rows of Table I. The average of the forward and backward calculations ($SG_{20,av}$) is presented in the next row. Furthermore, a SG simulation over 40 ps was carried out ($SG_{40}$). The relative errors of the nonaveraged results range up to 16%. Taking the average over forward and backward SG brings the relative errors below 10%.

***TI with short equilibration and sampling.*** In TI, a relatively low (e.g., 18) number of $\lambda$ values is used for which the system is thoroughly equilibrated (e.g., 50 ps) and sampled (e.g., 100 ps). SG is a limiting case of TI in which the equilibration and sampling periods tend to zero, and the number of $\lambda$ values tend to the number of timesteps in the simulation for which a $\partial H/\partial\lambda$ value was saved. A procedure that is intermediate between SG and TI is TI with short equilibration and sampling periods. The system is equilibrated for 2 ps at the starting value of $\lambda$ ($\lambda = 0$ or $\lambda = 1$, respectively), and then $\lambda$ is varied over the same 18 values as described in Free Energy Calculations for the Reference Compounds. The final structure at every $\lambda$ point serves as the starting structure for the next. Sampling windows of 1 ps ($TI_1$) and 5 ps ($TI_5$) per $\lambda$-point

**TABLE II. Free Energy Differences $\Delta F^{class}$ of Protonation (in kJ/mol)†**

| | Asp | Glu | Cter | 7Glu | 18Asp | 35Glu |
|---|---|---|---|---|---|---|
| TI | 206.9 | 212.8 | 239.4 | 211.0 | 210.2 | 174.1 |
| $TI_{el}$ | 243.2 | 250.5 | 273.1 | 258.3 | 235.9 | 199.3 |
| $\Delta TI_{el}$ | 17.5 | 17.7 | 14.1 | 22.4 | 12.2 | 14.5 |

†Calculated by TI (see Table I), as well as the free energy differences as calculated by TI while keeping the van der Waals parameters fixed as in the case of the protonated form of the residue ($TI_{el}$). The differences with respect to $\Delta F^{class}$ including the change in van der Waals interaction ($\Delta TI_{el}$) are given in percent.

**TABLE III. Free Energy Differences $\Delta F_{shift} = \Delta F^{class}_{prot} - \Delta F^{class}_{ref}$ (in kJ/mol) for Three Sites in HEWL†**

| Site | $\Delta F_{shift}$ | $\Delta F_{shift,el.}$ | Deviation |
|---|---|---|---|
| 7Glu | −1.8 | 7.8 | 533.3 |
| 18Asp | 3.3 | −7.3 | 321.2 |
| 35Glu | −38.7 | −51.2 | 32.3 |

†Calculated including ($\Delta F_{shift}$) and excluding ($\Delta F_{shift,el}$) the change in van der Waals parameters of the carboxylic group on protonation. The deviation of $\Delta F_{shift,el.}$ from $\Delta F_{shift}$ is given in percent.

were tested, always omitting the first 0.05 ps as equilibration time. The value of $\partial H/\partial\lambda$ was saved every 0.01 ps. For the 1-ps windows, the free energy difference for the reverse reaction path ($TI_{1,rev}$) and the average of the forward and backward reaction ($TI_{1,av}$) were calculated. The results can be seen in the last four rows of Table I. Although the single values of $TI_1$ and $TI_{1,rev}$ are not very accurate, their averages $TI_{1,av}$ are within a 5% relative error to the TI results. The results obtained by $TI_5$ are not much worse; however, they require much more computational effort.

Considering these results, $TI_{1,av}$ was chosen to calculate $\Delta F^{class}_{prot}$ for use in the MC procedure. Both pathways can be calculated simultaneously, starting from states $\lambda = 0$ and $\lambda = 1$, respectively.

### *Effect of changing the van der Waals interaction on protonation*

The standard GROMOS96 force field uses different van der Waals interaction parameters for the oxygens in the protonated and deprotonated state of the carboxylic acid groups. The values of the parameters can be found in Ref. 31. The contribution to the free energy of protonation of changing these van der Waals parameters on protonation was studied by calculating the free energy difference of protonation for the six test cases mentioned above through TI while keeping the van der Waals parameters fixed at the values for the residue in its protonated state. The results are shown in Table II.

Although the main contribution to $\Delta F^{class}$ is indeed caused by electrostatic interactions, the contribution coming from the change in van der Waals parameters on protonation is still 20%. Furthermore, it can be seen that the van der Waals contributions vary between different environments, which implies that they need not cancel when calculating the difference $\Delta F_{shift} = \Delta F_{prot} - \Delta F_{ref}$. To underscore this point, $\Delta F_{shift}$ as calculated including ($\Delta F_{shift}$) and excluding ($\Delta F_{shift,el}$) the van der Waals contribution is shown in Table III.

For two of the sites (7Glu and 18Asp), the sign of $\Delta F_{shift}$ changes when considering the van der Waals contribution, yielding relative errors of several 100%. Because the van der Waals parameters of ionizable sites of a molecule will in general be different in the different charge or protonation states, the van der Waals interaction will contribute to the free energy of protonation. Because this contribution seems to be sizeable using van der Waals parameters of commonly used biomolecular force fields, a calculation of the protonation free energy that is solely based on electrostatic interactions is likely to miss sizeable contributions to the free energy of protonation.

### Equilibration

Starting from the solvated X-ray structure of HEWL, a steepest descent energy minimization was carried out. The initial velocities were randomly chosen out of a Maxwell distribution. To further relax the system, 20 ps of MD simulation was run by applying positional restraints on all solute atoms. The force constant for the positional restraints was reduced from $2.5 \cdot 10^4$ to $2.5 \cdot 10^2$ to $2.5 \cdot 10^1$ to 0 kJ $mol^{-1}$ $nm^{-2}$ every 5 ps. Equally, for each of the solvated reference compounds, a steepest descent minimization was carried out. Because the number of conformational degrees of freedom of the reference compounds is very low, no initial MD simulation for equilibration was needed. The initial velocities were also chosen randomly out of Maxwell distributions.

As for the protonation state of the system, an initial equilibration phase was carried out for each of the three pH values considered. Both, starting the equilibration with all acidic residues completely protonated and completely deprotonated were tested. Both simulations converged after about 30 MC (de)protonation steps. The production runs were started from the relaxed crystal structure with the protonation states as indicated in Table IV.

### Generation of the Trajectories

Because the $TI_{1,av}$ procedure was used to calculate $\Delta F$ for the MC (de)protonation steps, two trajectories were generated at each MC step: one for the forward ($\lambda = 0 \rightarrow 1$) and one for the backward ($\lambda = 1 \rightarrow 0$) protonation path. There were no separate MD simulations added between the MC steps; therefore, the system was in a perturbed state ($0 < \lambda < 1$) for most of the simulation time. However, because only one ionizable site was perturbed at a time and because the same site remained unperturbed during a

**TABLE IV. Protonation State of the Acidic Residues After 100 MC Steps of Equilibration†**

| pH | 7 | 18 | 35 | 48 | 52 | 66 | 87 | 101 | 119 | 130 |
|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| 3 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| 2 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 |

†0 refers to the deprotonated and 1 to the protonated state of the corresponding residue indicated by its residue number in the HEWL amino acid sequence.

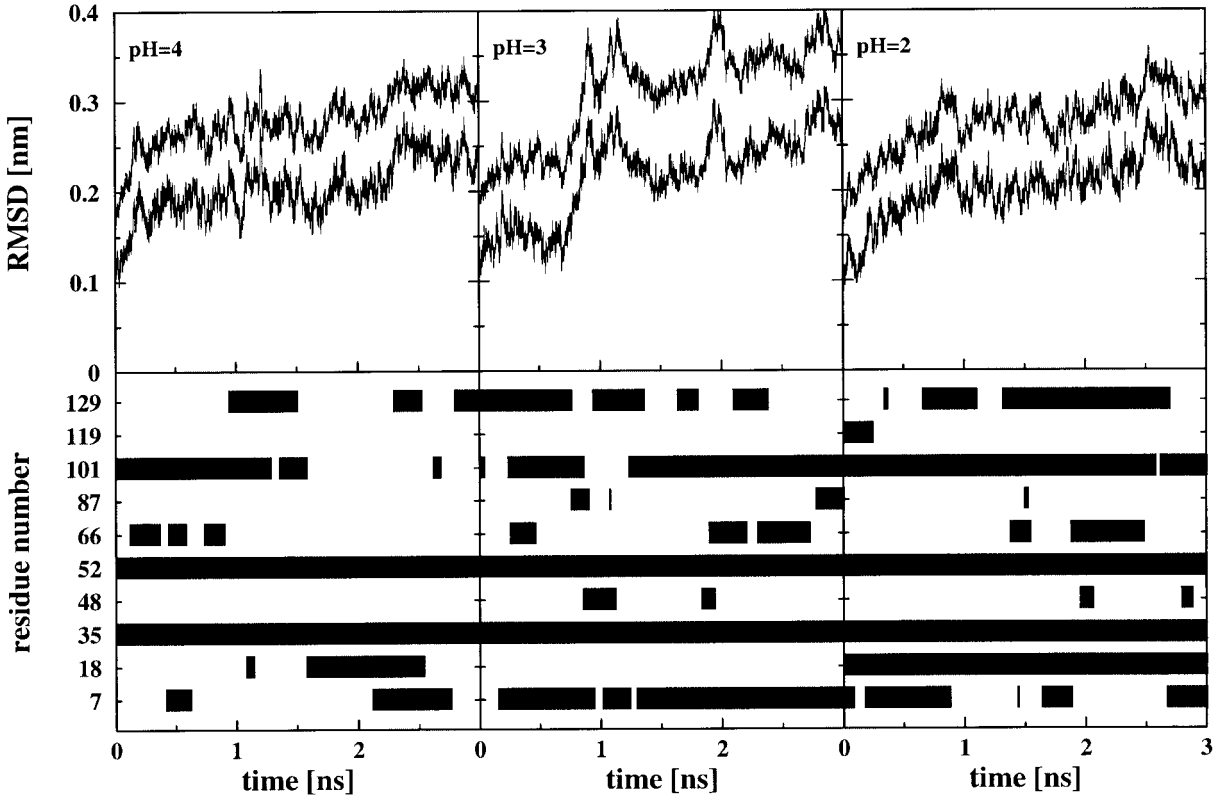


Fig. 1. All atom and backbone atom-positional root-mean-square deviation (RMSD) of the trajectory structures of hen egg white lysozyme (HEWL) with respect to the X-ray structure (upper panels) as well as the protonation state of the acidic residues (lower panels) as a function of time for the three constant pH simulations at pH = 4 (left panels), pH = 3 (middle panels) and pH = 2 (right panels). A black bar indicates that the residue is protonated.

number of other MC steps, this trajectory can be used for structure analysis. The generation of a continuous trajectory is possible when counting the trajectory along the forward reaction path in case of acceptance and along the backward reaction path in case of rejection of the MC step.

## RESULTS AND DISCUSSION

As mentioned in Theory and Methods, three simulations were carried out, at pH 4, 3, and 2. All simulations are 3 ns long. Figure 1 shows the all atom and backbone atom-positional root-mean-square deviation (RMSD) of the three trajectories from the X-ray structure (upper panels) as well as the protonation state of the acidic residues (lower panels) as a function of time.

Lysozyme is a rather stable protein; therefore, no major structural changes are expected, even at a low pH. However, when comparing Figure 1 with the simulation of HEWL in solution at pH 6 of Stocker et al.,[45,46] we note that structural changes occur during the simulations at low pH that do not occur at pH 6; the RMSD values of all three simulations at low pH are higher than observed at pH 6 (0.14 nm for $C_\alpha$ atoms and 0.21 nm for all atoms). Furthermore, the simulation at pH 3 (middle panels) shows that structural changes are correlated with changes in the protonation states: as soon as 48Asp or 87Asp are protonated for a longer period of time, structural changes are induced. We investigate these changes later in more detail.

From Figure 1, we can also get an impression of the timescale of protonation and deprotonation events in proteins. The lifetimes of the protonation states of a residue vary considerably from tens of picoseconds to

**TABLE V. Experimental Apparent $pK_A$ Values† for Hen Lysozyme[5,6]**

| | Experiment | | Simulation | | |
|---|---|---|---|---|---|
| Residue | HEWL | Ref | pH = 4 | pH = 3 | pH = 2 |
| 7Glu | 2.85 | 4.4 | 3.6 (0.8) | 5.0 (1.4) | 1.1 (1.6) |
| 18Asp | 2.66 | 4.0 | 3.5 (2.2) | 1.2 (0.9) | 5.1 (1.3) |
| 35Glu | 6.20 | 4.4 | 11.9 (2.3) | 11.7 (2.5) | 13.3 (3.5) |
| 48Asp | <2.5 | 4.0 | −0.5 (1.2) | 0.5 (2.4) | −0.0 (1.7) |
| 52Asp | 3.68 | 4.0 | 8.0 (1.3) | 7.6 (1.0) | 10.5 (1.9) |
| 66Asp | <2.0 | 4.0 | 0.7 (3.4) | 2.6 (1.5) | 1.2 (3.0) |
| 87Asp | 2.07 | 4.0 | 0.7 (1.3) | 0.5 (2.3) | −1.7 (2.1) |
| 101Asp | 4.09 | 4.0 | 4.7 (1.1) | 5.1 (1.8) | 4.2 (1.3) |
| 119Asp | 3.20 | 4.0 | 0.1 (1.3) | −0.9 (1.2) | −0.6 (1.1) |
| 130Cter | 2.75 | 3.8 | 4.0 (1.6) | 2.9 (1.6) | 2.9 (1.7) |

†Reference $pK_A$ values for single amino acids in solution,[28] and effective average $pK_A$ values (with standard deviations within parentheses) for the acidic residues of lysozyme as calculated by using Eqs. 1 and 5 from the three constant pH simulations at pH 4 3, and 2.

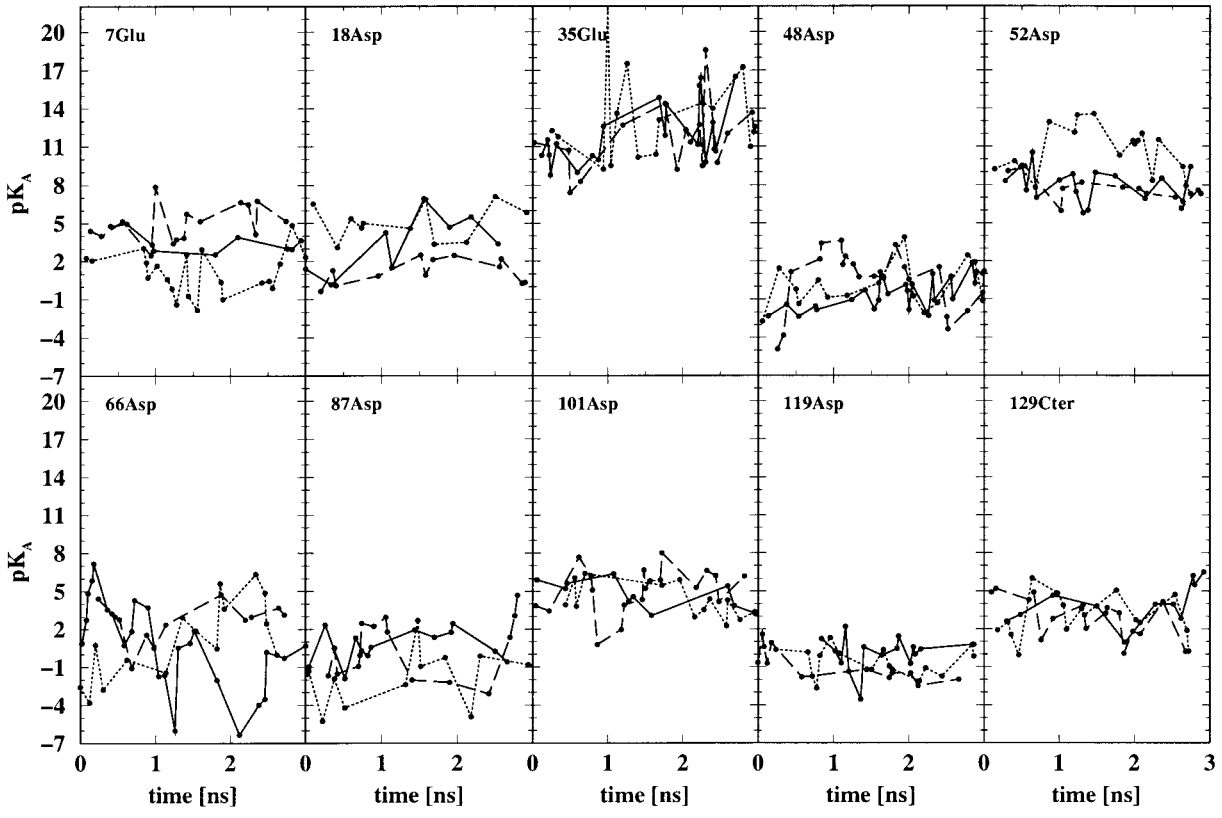


Fig. 2. Effective $pK_A$ values for the acidic residues of hen lysozyme as a function of time as calculated by using Eqs. 1 and 5 for the three constant pH simulations at pH 4, 3, and 2. The solid lines indicate the results for pH 4, the dashed lines the ones for pH 3, and the dotted lines the ones for pH 2.

nanoseconds. Because of the insufficient statistics, it is hard to draw conclusions in terms of the population of the protonation states at different pH. Considering, for instance, 18Asp, we can see that it is protonated for about 100 ps and 1 ns at pH 4 (lower left panel), not protonated at all at pH 3 (lower middle panel), and protonated during the entire simulation at pH 2 (lower right panel). From this we can conclude that the protein samples a different region of the phase space in the three simulations. Although 18Asp is always in conformations disfavoring protonation at pH 3, there is an equilibrium of conformations favorable and unfavorable to protonation at pH 4.

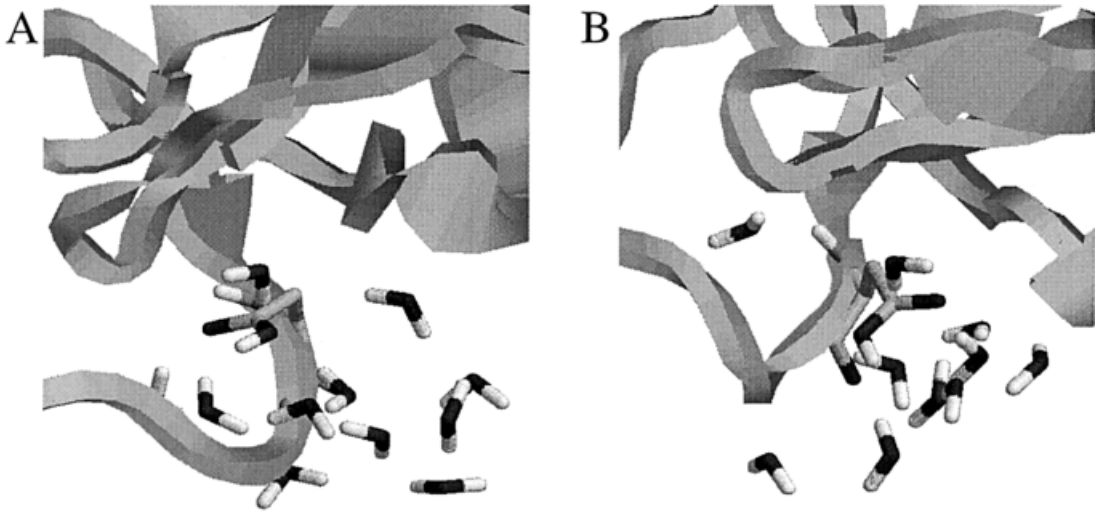


Fig. 3. Snapshots (A at 120 ps, B at 1860 ps) of residue 66Asp and its environment from the 3-ns constant pH simulation of hen lysozyme at pH = 2. Hydrogen atoms of acidic residues are drawn irrespective of the actual protonation state (given in Fig. 1) of a residue. All water molecules within a distance of 0.4 nm from atoms of 66Asp are drawn.

However, there are residues that seem better equilibrated. Considering 101Asp (experimental $pK_A = 4.09$), we can see that it is protonated for about 50% of the time at pH 4, for about 82% at pH 3, and for about 99% of the time at pH 2. Therefore, 101Asp does not seem to interact strongly with other titratable sites—its protonation behavior is the one that is expected for a single amino acid.

By changing the pH from 4 to 3, the average protonation increases for six residues, it decreases for 18Asp, whereas 35Glu and 52Asp stay completely protonated and 119Asp stays completely deprotonated. Changing the pH from 3 to 2, four residues become more protonated, two residues do not change their protonation state, and four residues become slightly less protonated on average. If there were no interaction between titratable residues and the environment of the residues did not change as a function of pH, the Metropolis criterion Eq. (6) would ensure that protonation increases when lowering the pH. In this case, lowering the pH of the system by one unit would cause the residue to be protonated by a factor of 10 more. The fact that this is not the case in Figure 1 indicates a strong coupling between residues, for which the titration curves are distorted.[12]

To investigate the quality of the algorithm presented in this work we compare the $pK_A$ values of the acidic residues as calculated from the simulations with their experimental values. When referring to $pK_A$ values in this work, we always refer to the so-called effective $pK_A$ defined by Eq. 1. Note that this does not correspond to the so-called apparent $pK_A$, defined as the pH at which a residue is 50% protonated and 50% deprotonated; the $pK_A$ as defined by Eq. 1 can be considered a pH and time-dependent quantity. Because the $pK_A$ values derived from experiment are apparent $pK_A$ values, the comparison of the calculated with the experimental $pK_A$ values is not straightforward. The closer the apparent $pK_A$ is to the simulated pH, the less the apparent $pK_A$ should deviate from the effective $pK_A$. However, if the apparent $pK_A$ is far away from the simulated pH, the difference between the apparent and explicit $pK_A$ values can be large. Another deviation of the simulation results from experiment is introduced by the lack of salt ions in the simulation setup (see also Counter Ions).

The effective $pK_A$ values as calculated from the MD trajectories using Eqs. 1 and 5 are listed as ensemble averages in Table V and plotted as a function of time in Figure 2.

The results presented in Table V are in qualitative agreement with experiment: the $pK_A$ shifts with respect to the reference $pK_A$ values generally point in the right direction, except for 52Asp. The magnitude of the shifts is rather too large. However, for 7Glu, 18Asp and 66Asp, the experimental value lies in between the results obtained from the three simulations, and for 48Asp, 52Asp, 87Asp, 101Asp, and 129Cter, there may be a slight trend toward the experimental value discernible in Figure 2. Figure 2 and Table V show that much better statistics are needed to obtain accurate $pK_A$ values. The 3-ns simulation time was not sufficient to completely relax the structure of the protein. Furthermore, it has been shown that the $pK_A$ values, as calculated in MC simulations of the protonation state of a protein, converge rather slowly, even if the structure of the protein does not change significantly.[47,48] In agreement with these observations, the $pK_A$ values fluctuate over several $pK_A$ units as a function of time, with fluctuation periods ranging from tens of picoseconds to nanoseconds. This illustrates the inhomogeneity in time and space of the environment of residues leading to a distortion of the titration curves. The calculated $pK_A$ of 35Glu is very large. However, because the experimental apparent $pK_A$ is 2.2 units higher than the highest pH simulated, the effective $pK_A$ can differ significantly from the apparent one. To compare the $pK_A$ of 35Glu, as obtained through simulation with the one obtained through experiment, simulations at a pH corresponding roughly to

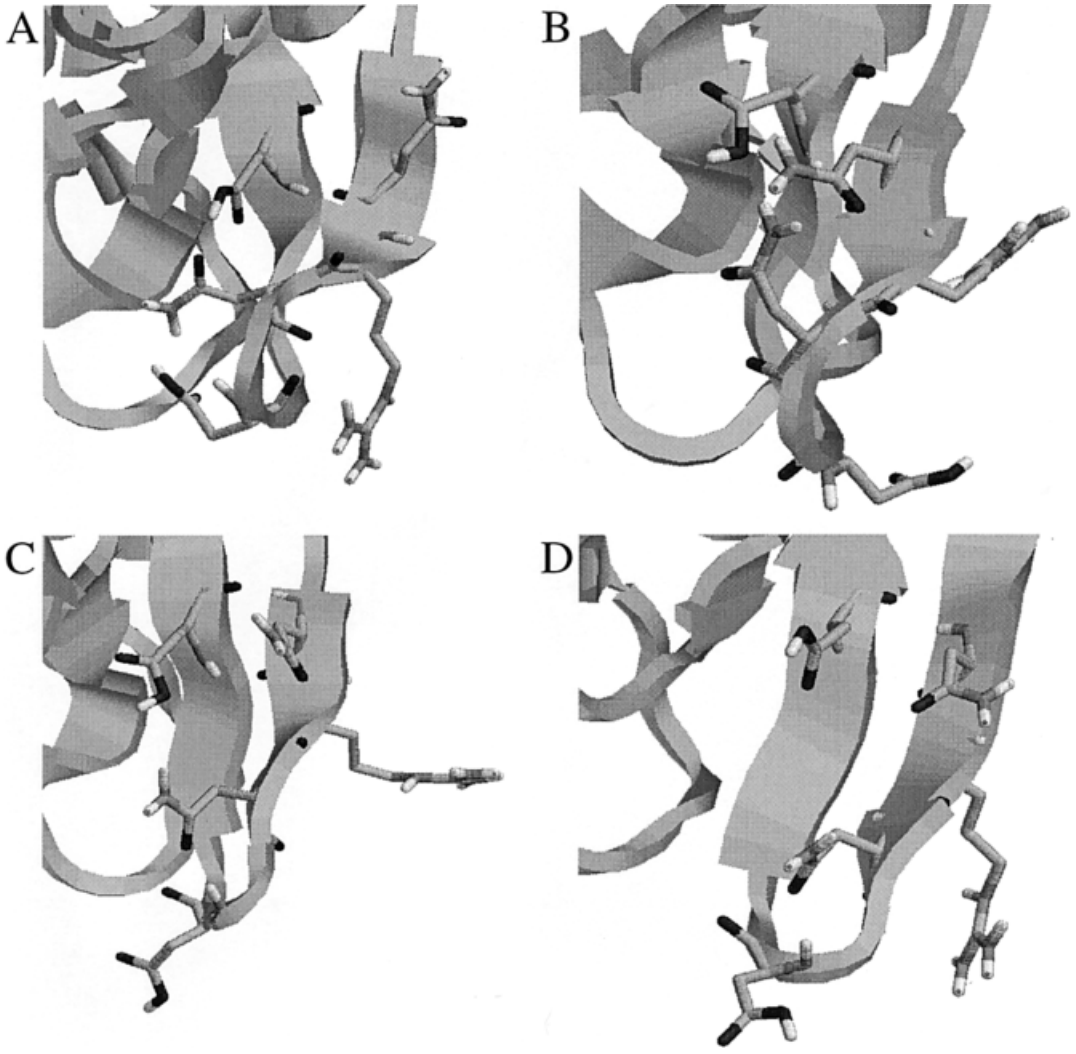


Fig. 4. Snapshots (A at 1460 ps, B at 2740 ps) of the β-sheet 42Ala—60Ser from the 3-ns constant pH simulation of hen lysozyme at pH = 2. Snapshots C (at 1820 ps) and D (at 1940 ps) are from the simulation at pH = 3. The residues drawn are (clockwise starting at the upper right part) 44Asn, 45Arg, 48Asp, 46Asn, and 52Asp. See also caption of Figure 3.

the apparent $pK_A$ (i.e., simulations at a pH around 6) would have to be performed. At a significantly higher pH, it would be necessary to include other titratable residues with a low $pK_A$, such as 15His, into the MC steps.

A comparison of the fluctuations shown in different panels of Figure 2 in more detail shows that the protonation state of the other residues is not the only cause for changing the $pK_A$ value of a particular residue. In general, the $pK_A$ value of a residue fluctuates even when the protonation states of the other residues do not change. The fluctuations in the structure of the protein seem to be an important cause for fluctuations of the $pK_A$ values. We may classify three different types of structural fluctuations. The first type are side-chain fluctuations. The $pK_A$ of a residue may change significantly if the side-chain moves into or out of the protein, or if another charged side-chain moves closer or farther away from the residue under consideration. A side-chain fluctuation of the former type is illustrated in Figure 3 for 66Asp.

Two snapshots of the simulation at pH 2 are shown. We note that in all figures with protein structures, the hydrogen atoms of the acidic residues are drawn irrespective of the protonation state of the residue. In case the residue is not protonated, the drawn hydrogen atom is in the position of the dummy atom. Figure 3A shows the side-chain of 66Asp after 120 ps of simulation pointing into the protein, and Figure 3B after 1860 ps of simulation pointing out of the protein into the solvent. All water molecules within a radius of 0.4 nm around the atoms of 66Asp are drawn. The $pK_A$ value of 66Asp for the structure shown in Figure 3A is −3.8; the one for the structure in Figure 3B is 5.6. See also Figure 2.

The second type of structural fluctuations are fluctuations of the secondary structure. Depending on the protonation state of certain residues, the secondary structure can adapt its shape slightly to the given charge state. For example, we consider the β-sheet 42Ala—60Ser which contains the ionizable residues 48Asp and 52Asp.

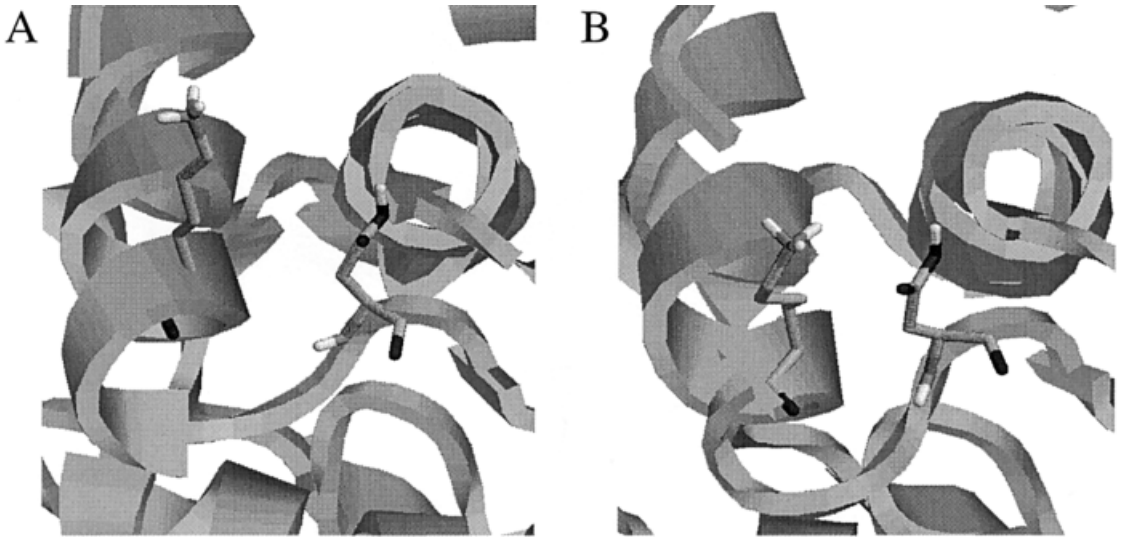


Fig. 5. Snapshots (A at 340 ps, B at 1560 ps) of residues 13Lys and 18Asp and their environment from the 3-ns constant pH simulation of hen lysozyme at pH = 4. See also caption of Figure 3.

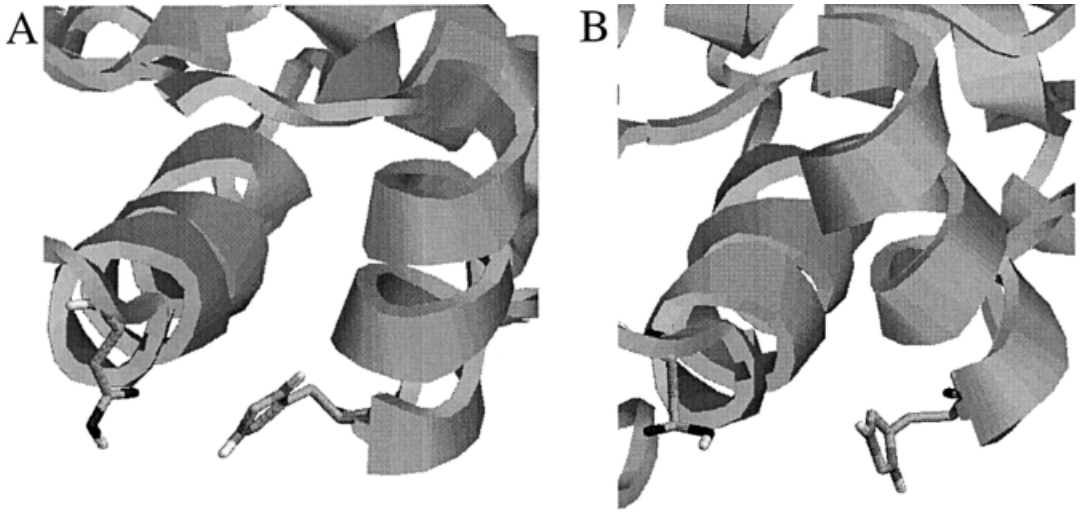


Fig. 6. Snapshots (A the crystal structure, B after 2800-ps simulation at pH = 3) of residues 15His and 87Asp and their environment in hen lysozyme. See also caption of Figure 3.

Figure 4A shows the β-sheet at pH 2 after 1460 ps. The structure of the β-sheet in Figure 4A is very similar to the one in the crystal structure. The side-chain of 48Asp points into the protein, and the one of 46Asn is caught between 48Asp and 52Asp. The $pK_A$ of 52Asp is thus influenced by the negative partial charge of the oxygen atom of 46Asn pointing toward the hydrogen of 52Asp, resulting in a $pK_A$ of 13.5 for 52Asp. Later in this simulation, at 2740 ps (Fig. 4B), the side-chain of 48Asp is sticking out of the protein, straightening the β-sheet slightly. Instead of the oxygen of 46Asn, the hydrogens of 44Asn and 46Asn with a positive partial charge point to the oxygen of 52Asp, lowering the $pK_A$ of this residue to 7.3. The $pK_A$ of 48Asp, on the other hand, changes from around 0 in the structure in Figure 4A to around 2.5 for the structure shown in Figure 4B.

The simulation at pH 3 shows that the structural change of the β-sheet can be even more drastic. Figure 4C shows the structure of the β-sheet at pH 3 after 1820 ps of simulation. The side-chain of 48Asp is sticking out of the protein. However, although 48Asp gets protonated for a longer period of time at around 900 ps, the β-sheet is completely flat. In this structure of the β-sheet, the $pK_A$ of 48Asp is chiefly dependent on the position of the side-chain of 45Arg. At 1820 ps, the side-chain of 45Arg points away from 48Asp (Fig. 4C), the $pK_A$ of 48Asp is 3.3, and the residue becomes protonated (see Fig. 1). At 1940 ps in the same simulation (Fig. 4D), the side-chain of 45Arg points toward 48Asp, the $pK_A$ of 48Asp shifts down to 1.5, and the residue becomes deprotonated.

The third type of fluctuations are fluctuations of the tertiary structure. Figure 5 shows 13Lys and 18Asp in the simulation at pH 4 after 340 ps (Fig. 5A) and 1560 ps (Fig. 5B) of simulation time. The residues are not part of the same secondary structure element; 13Lys is part of a helix

and 18Asp is part of a loop. Depending on the tertiary structure of the protein, these two residues can be closer (0.45 nm in Fig. 5B) or further apart (0.68 nm in Fig. 5A). Accordingly, the $pK_A$ of 18Asp changes from 0.2 (340 ps, Fig. 5A) to 6.9 (1560 ps, Fig. 5B).

The effect of tertiary structure fluctuations is even more obvious in the example of 87Asp. Figure 6 shows the environment of the residues 15His and 87Asp in the crystal structure (Fig. 6A) and after 2800 ps of simulation at pH 3 (Fig. 6B). The two residues are located at the ends of two neighboring helices (residues 4Gly–14Arg and residues 88Ile–99Val). As can be seen, the helices change their position with respect to each other, so that the separation between 15His and 87Asp is increased from 0.6 nm in the crystal structure to 0.86 nm after 2800 ps of simulation at pH 3. The $pK_A$ of 87Asp changes from around $-1$ to 4.7, depending on its relative position to the protonated side chain of 15His.

All three types of structural fluctuations contribute to the $pK_A$ fluctuations of the acidic residues shown in Figure 2. The side-chain fluctuations are the fastest ones resulting in the higher frequency fluctuations. The secondary and tertiary structure fluctuations induce lower frequency modes that are not well sampled in the 3 ns simulated.

## CONCLUSIONS

We have presented an algorithm to simulate proteins at constant pH combining molecular dynamics (MD) and Monte Carlo (MC) simulation. The free energies of protonating or deprotonating a residue are calculated through MD simulation, whereas MC steps sample the protonation states of the ionisable residues during an MD simulation. Although one residue is always perturbed during the MD simulation, the trajectory generated during the free energy calculations can be used for structural analysis, because the perturbation of a specific residue occurs only during an attempt to change its protonation state and the residue is not perturbed otherwise. Thus, the algorithm presented makes it possible to generate trajectories at a Boltzmann distributed ensemble of protonation states. In addition, the solvent degrees of freedom are explicitly simulated. This makes the algorithm suitable for both, pH-dependent structural studies and detailed investigations of titration properties. The algorithm for constant pH simulations considers the neutral and charged form of residues explicitly rather than representing them by intermediate or extrapolated charge states.

The algorithm was tested by application to hen egg white lysozyme (HEWL) in aqueous solution and comparing the $pK_A$ values of the acidic residues as calculated from the constant pH simulations with the experimentally measured values. Although 3 ns of simulation are too short for the $pK_A$ values to be converged, the values calculated are in qualitative agreement with the experimental data.

The present study indicates that the environment of the titratable residues influences the $pK_A$ values to a large extent and that the $pK_A$ values fluctuate as a function of time over several $pK_A$ units as a result of local structural changes. Three types of structural fluctuations were distinguished: fluctuations of side-chains, of the secondary structure, and of the tertiary structure. They cover a time range of hundreds of picoseconds to nanoseconds. The slowest structural fluctuations are insufficiently sampled when only averaging over subnanosecond trajectories. Furthermore, several structural changes only occur when the protonation states of the residues change. Therefore, it is essential to allow for these changes in protonation state to occur in the simulations to have properly weighted structures in the trajectory.

## ACKNOWLEDGMENTS

R.B. thanks Kim Sharp for helpful discussions on $pK_A$ calculations.

## REFERENCES

1. Stoeckelhuber M, Noegel AA, Eckerskorn C, Kohler J, Rieger D, Schleicher M. Structure/function studies on the pH-dependent actin-binding protein hisactophilin in dictyostelium mutants. J Cell Sci 1996;109:1825–1835.
2. Dimitrov RA, Crichton RR. Self-consistent field approach to protein structure and stability. I. pH dependence of electrostatic contribution. Proteins 1997;27:576–596.
3. Ruano MLF, Perez-Gil J, Casals C. Effect of acidic pH on the structure and lipid binding properties of porcine surfactant protein a—potential role of acidification among its exocytic pathway. J Biol Chem 1998;273:15183–15191.
4. Renugopalakrishnan V, Dobbs JC, Collette TW, Carreira LA, Hutson TB, Garduno-Juarez R. Human pancreatic thread protein, an exocrine thread protein with possible implications to Alzheimer's disease: secondary structure in solution at acid pH. Biochem Biophys Res Commun 1999;258:653–656.
5. Kuramitsu S, Hamaguchi K. Analysis of the acid-base titration curve of hen lysozyme. J Biochem 1980;87:1215–1219.
6. Bartik K, Redfield C, Dobson CM. Measurement of the individual $pK_A$ values of acidic residues of hen and turkey lysozymes by two-dimensional $^1$H NMR. Biophys J 1994;66:1180–1184.
7. de Prat-Gay G. Determination of the $pK_A$ of a glutamic residue in a protein by difference titration capillary electrophoresis. Protein Peptide Lett 1998;5:9–14.
8. Mertz JE, Pettitt BM. Molecular dynamics at a constant pH. Supercomputer Applications and High Performance Computing 1994;8:47–53.
9. Baptista AM, Martel PJ, Petersen SB. Simulation of protein conformational freedom as a function of pH: constant-pH molecular dynamics using implicit titration. Proteins 1997;27:523–544.
10. Börjesson U, Hünenberger PH. Explicit-solvent molecular dynamics simulation at constant pH: methodology and application to small amines. J Chem Phys 2001;114:9706–9719.
11. Onufriev A, Case DA, Ullmann GM. A novel view of pH titration in biomolecules. Biochemistry 2001;40:3413–3419.
12. Yang A, Gunner MR, Sampogna R, Sharp K, Honig B. On the calculation of $pK_A$s in proteins. Proteins 1993;15:252–265.
13. Atkins PW. Physical chemistry. Oxford: Oxford University Press; 1999.
14. Linderstrøm-Lang K. On the ionisation of proteins. C. R. Trav. Lab. Carlsberg 1924;15:1–29.
15. Zauhar RJ, Morgan RS. A new method for computing the macromolecular electric potential. J Mol Biol 1985;186:815–820.
16. Lim C, Bashford D, Karplus M. Absolute $pK_A$ calculations with continuum dielectric methods. J Phys Chem 1991;95:5610–5620.
17. Sitkoff D, Sharp KA, Honig B. Accurate calculation of hydration free energies using macroscopic solvent models. J Phys Chem 1994;98:1978–1988.
18. States DJ, Karplus M. A model for electrostatic effects in proteins. J Mol Biol 1987;197:122–130.
19. Antosiewicz J, McCammon JA, Gilson MK. Prediction of pH-dependent properties of proteins. J Mol Biol 1994;238:415–436.
20. Demchuk E, Wade RC. Improving the continuum dielectric approach to calculating $pK_A$s of ionizable groups in proteins. J Phys Chem 1996;100:17373–17387.
21. Schaefer M, Sommer M, Karplus M. pH-dependence of protein

stability: absolute electrostatic free energy differences between conformations. J Phys Chem 1997;101:1663–1683.
22. Gibas CJ, Subramaniam S. Explicit solvent models in protein $pK_A$ calculations. Biophys J 1996;71:138–147.
23. van Vlijmen HWT, Schaefer M, Karplus M. Improving the accuracy of protein $pK_A$ calculations: conformational averaging versus the average structure. Proteins 1998;33:145–158.
24. Russel ST, Warshel A. Calculations of electrostatic energies in proteins. The energetics of ionized groups in bovine pancreatic trypsin inhibitor. J Mol Biol 1985;185:389–404.
25. Sham YY, Chu ZT, Warshel A. Consistent calculations of pK(a)'s of ionizable residues in proteins: semi-microscopic and microscopic approaches. J Phys Chem B 1997;101:4458–4472.
26. Del Buono GS, Figueirido FE, Levy RM. Intrinsic $pK_A$s of ionizable residues in proteins: an explicit solvent calculation for lysozyme. Proteins 1994;20:85–97.
27. Schutz CN, Warshel A. What are the dielectric "constants" of proteins and how to validate electrostatic models? Proteins 2001; 44:400–417.
28. Fersht A. Enzyme structure and mechanism. New York: W.H. Freeman and Company; 1985. p 155–175.
29. van Gunsteren WF, Billeter SR, Eising AA, Hünenberger PH, Krüger P, Mark AE, Scott WRP, Tironi IG. Biomolecular simulation: the GROMOS96 manual and user guide. Zürich: Vdf Hochschulverlag AG an der ETH Zürich; 1996.
30. Scott WRP, Hünenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Krüger P, van Gunsteren WF. The GROMOS biomolecular simulation program package. J Phys Chem A 1999;103:3596–3607.
31. Bürgi R, Läng F, van Gunsteren WF. A comparison of seven fast but approximate methods to compute the free energy of deprotonation for amino acids. Mol Sim 2001;27:215–236.
32. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In: Pullman B, editor. Intermolecular forces. Dordrecht, The Netherlands: Reidel; 1981. p 331–342.
33. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
34. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 1977;23:327–341.
35. Tironi IG, Sperb R, Smith PE, van Gunsteren WF. A generalized reaction field method for molecular dynamics simulations. J Chem Phys 1995;102:5451–5459.
36. Smith PE, van Gunsteren WF. Reaction field effects on the simulated properties of liquid water. Mol Simul 1995;15:233–245.
37. Ibragimova GT, Wade RC. Importance of explicit salt ions for protein stability in molecular dynamics simulation. Biophys J 1998;74:2906–2911.
38. Chowdhuri S, Chandra A. Molecular dynamics simulations of aqueous NaCl and KCl solutions: effects of ion concentration on the single-particle, pair, and collective dynamical properties of ions and water molecules. J Chem Phys 2001;115:3732–3741.
39. Beutler TC, Mark AE, van Schaik RC, Gerber PR, van Gunsteren WF. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. Chem Phys Lett 1994;22:529–539.
40. Kirkwood JG. Statistical mechanics of fluid mixtures. J Chem Phys 1935;3:300–313.
41. van Gunsteren WF, Beutler TC, Fraternali F, King PM, Mark AE, Smith PE. Computation of free energy in practice: choice of approximations and accuracy limiting factors. In: van Gunsteren WF, Weiner PK, Wilkinson AJ, editors. Computer simulation of biomolecular systems, theoretical and experimental applications. Vol. 2. Leiden, The Netherlands: ESCOM Science Publisher; 1993. p 315–348.
42. Berendsen HJC, Postma JPM, van Gunsteren WF. Statistical mechanics and molecular dynamics: the calculation of free energy. In: Hermans J, editor. Molecular dynamics and protein structure. Western Springs: Polycrystal Book Service; 1985. p 43–46.
43. van Gunsteren WF. Methods for calculation of free energies and binding constants: successes and problems. In: van Gunsteren WF, Weiner PK, editors. Computer simulation of biomolecular systems, theoretical and experimental applications. Leiden, The Netherlands: Escom Science Publishers; 1989. p 27–59.
44. Yun-yu S, Mark AE, Cun-xin W, Fuhua H, Berendsen HJC, van Gunsteren WF. Can the stability of protein mutants be predicted by free energy calculations? Protein Eng 1993;6:289–295.
45. Stocker U, van Gunsteren WF. Molecular dynamics simulation of hen egg white lysozyme: a test of the GROMOS96 force field against nuclear magnetic resonance data. Proteins 2000;40:145–153.
46. Stocker U, Spiegel K, van Gunsteren WF. On the similarity of properties in solution or crystalline state: a molecular dynamics study of hen lysozyme. J Biomol NMR 2000;18:1–12.
47. Beroza P, Case DA. Calculations of proton-binding thermodynamics in proteins. Methods Enzymol 1998;295:170–189.
48. Spassov VZ, Bashford D. Multiple-site ligand binding to flexible macromolecules: separation of global and local conformational change and an iterative mobile clustering approach. J Comput Chem 1999;20:1091–1111.