Calculation of Protein Ionization Equilibria With Conformational Sampling: pK_a of a Model Leucine Zipper, GCN4 and Barnase

Alemayehu A. Gorfe, Philippe Ferrara, Amedeo Caflisch, Daniel N. Marti, Hans Rudolf Bosshard, and Ilian Jelesarov*

Department of Biochemistry, University of Zurich, Zurich, Switzerland

ABSTRACT **The use of conformational ensembles provided by nuclear magnetic resonance (NMR) experiments or generated by molecular dynamics (MD) simulations has been regarded as a useful approach to account for protein motions in the context of pK_a calculations, yet the idea has been tested occasionally. This is the first report of systematic comparison of pK_a estimates computed from long multiple MD simulations and NMR ensembles. As model systems, a synthetic leucine zipper, the naturally occurring coiled coil GCN4, and barnase were used. A variety of conformational averaging and titration curve-averaging techniques, or combination thereof, was adopted and/or modified to investigate the effect of extensive global conformational sampling on the accuracy of pK_a calculations. Clustering of coordinates is proposed as an approach to reduce the vast diversity of MD ensembles to a few structures representative of the average electrostatic properties of the system in solution. Remarkable improvement of the accuracy of pK_a predictions was achieved by the use of multiple MD simulations. By using multiple trajectories the absolute error in pK_a predictions for the model leucine zipper was reduced to as low as approximately 0.25 pK_a units. The validity, advantages, and limitations of explicit conformational sampling by MD, compared with the use of an average structure and a high internal protein dielectric value as means to improve the accuracy of pK_a calculations, are discussed. Proteins 2002;46:41–60.**

Key words: coiled coil; NMR; molecular dynamics; ensemble; salt bridge; cluster; electrostatics

INTRODUCTION

Release and uptake of protons are fundamental to a variety of biological processes such as enzymatic catalysis, macromolecular stability, and formation of macromolecular complexes. Determination of the pK_a values of protein ionizable groups is key to the understanding of the pH-dependent properties of proteins. Theoretical studies of the ionization behavior of protein charged groups are a vital part of this endeavor. Structure-based predictions of the electrostatic properties of proteins have motivated researchers ever since the beginning of modern structural biology,[1,2] and considerable advances have been made in the last two decades.[3–8] Quantitative computational predictions of pK_a values are still problematic in many cases.[6,9]

The titration behavior of proteins is governed by the ionization equilibria of their acidic and basic charged groups. The pK_a of a group in a folded conformation of a protein can deviate by several pH units from the pK_a of a fully solvated model compound because the electrostatic environments experienced by a group in the protein and in solution are very different. Interactions of charged ionizable groups with the charged form of other titratable groups (site-site interactions) and with permanent dipoles in the protein (background interactions), as well as the altered interactions with water (Born energy or desolvation energy) are the main sources of pK_a shifts.[10] The change of the ionization equilibrium of charged groups and the titration behavior of proteins are commonly calculated by the numerical solution of the linearized form of the Poisson Boltzmann (FDPB) equation for a group in a protein relative to an isolated model compound in solution.[4] In the continuum representation, a dielectric boundary between the bulk solvent and the protein is defined, usually represented by the solvent accessible surface of the molecule. The solvent is modeled as a continuum of high polarizability, whereas a lower dielectric constant is assigned to the protein domain, and the protein structure is treated at atomic level.[4–6] The pH-dependent free energies are calculated by Boltzmann summation over all important protonation states in a certain pH interval. As a result, the average degree of protonation of each ionizable group as a function of pH is obtained. This methodology was first applied to titration studies of lysozyme a decade ago[4] and has since been applied to many proteins.[11–13]

The Supplementary Material referred to in this section can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/46_1/v46_1.html

Grant sponsor: Bundesamt für Bildung und Wissenschaft; Grant number: 97.0592; Grant sponsor: Swiss National Science Foundation; Grant number: 31.55308.98.

*Correspondence to: Ilian Jelesarov, Department of Biochemistry, University of Zurich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland. E-mail: iljel@bioc.unizh.ch

Received 30 March 2001; Accepted 21 August 2001

DOI 10.1002/prot.10027

In the framework of the continuum method, the accuracy of pK_a calculation critically depends on several factors, such as the choice of partial atomic charges, atomic radii, and protein interior dielectric constant. The exact numerical value of the latter, for example, has been widely and critically debated.[14–16] Furthermore, the continuum method works on a rigid body approximation of the protein, the calculations being performed with fixed atomic positions, as seen in an X-ray structure or in a nuclear magnetic resonance (NMR) conformer. Therefore, it cannot accurately account for the existence of tautomeric states arising from the uncertainty in the location of the protons for such groups as protonated Glu or Asp and unprotonated His or for the conformational fluctuations of the molecule on local and global scales. In addition, the (mostly local) conformational change induced by the change of the electrostatic field on protonation and deprotonation of titratable groups cannot be captured by a rigid body treatment of a structure that does not change with pH.[17] The mentioned limitations have prompted for extensive theoretical work and many elegant solutions have been proposed over the years.[18–20]

Protein flexibility may be divided into two contributions,[17] namely, the fluctuation of the protein backbone and side chains resulting from thermal motion or conformational disorder, and local side chain motions induced by the titration per se. The effect of local structural fluctuations on the accuracy of pK_a predictions has been addressed by several authors.[18–20] Here we focus on the effect of global structural variability on pK_a predictions. Global conformational variability is suggested to have substantial impact on the titration behavior of proteins.[17] It may have to be explicitly accounted for if we want to properly simulate the electrostatic properties for the system in solution, which is best represented by an ensemble of conformational states in a dynamic equilibrium. Indeed, the use of an ensemble of structures from NMR experiments,[21,22] Monte Carlo (MC),[23] and Molecular Dynamics (MD) simulations[9,12,24] has improved the accuracy of pK_a predictions, compared with calculations on a single X-ray structure. Although NMR ensembles are derived from experimental interproton distances (and dihedral angles) and, hence, are likely to accurately represent the solution structure, they are biased toward the pH of the NMR experiment. MD simulations, on the other hand, are also biased toward the presumed protonation states of charged groups at the pH of the simulation. However, MD simulations may sample thermally accessible and thus statistically important conformations.[17] The sampling efficiency of MD calculations can be improved significantly by running multiple and/or long simulations. A survey of the literature reveals that pK_a calculations have been performed on ensembles derived from relatively short MD simulations (100–200 ps). It has been suggested that the use of more conformational sampling (longer simulation times and/or multiple simulations) might result in better predictions.[9] Surprisingly, to the best of our knowledge, that there is no published report dealing with a systematic comparison of pK_a estimates computed from NMR ensembles and MD simulations.

In this report, we present pK_a values for an engineered leucine zipper (henceforth referred to as AB zipper) calculated on ensembles of conformations derived either experimentally by NMR or from four relatively long (1 ns) MD simulations in explicit solvent. Different averaging schemes designed to account for the conformational variability, yet keeping the computational costs low, are considered and critically evaluated. The calculated pK_a values are tested against very accurate experimental pK_a values for 10 Glu residues published recently.[25] Although most of the present study is concerned with the synthetic leucine zipper designed in our laboratory, two other proteins are used to generalize the main arguments. These are the coiled coil domain GCN4p1, which is structurally similar to the AB zipper, and barnase. For the former, pK_a values are calculated by using the X-ray structure and ensembles from MD simulations and compared with available experimental pK_a values. For barnase, pK_a values are obtained by using several X-ray structures, NMR conformers, and a 700-ps MD simulation. We introduce, for the first time in the context of pK_a calculations, averaging protocols based on (population weighted) clusters of MD conformations. The advantages and limitations of explicit conformational sampling by MD for improving the accuracy of pK_a predictions are discussed.

MATERIALS AND METHODS

NMR and X-ray Structures

The solution structure of a synthetic leucine zipper, referred to as AB zipper in the following, has been determined (pdb entry 1FMH[25]). Calculations with the dimerization leucine zipper domain, GCN4p1, were performed by using the X-ray structure (pdb entry 2ZTA[26]). Several published crystal structures of barnase (EC 3.1.27.3), the small ribonuclease from *Bacillus amyloliquefaciens* were used for pK_a calculations (pdb entry codes 1RNB, 1B20,1BNI, 1B2X, 1B2Z[27,28]). In addition, pK_a values were computed on 20 NMR conformers of barnase (pdb entry 1BNR[29]).

MD Simulations

The MD simulations for the leucine zippers were carried out with the CHARMM program[30] by using the all atom parameter set of CHARMM22.[31] Hydrogen atoms were added to the coordinates of the X-ray structure of GCN4p1 by the HBUILD option[32] of CHARMM. The standard protonation states of the acidic and basic groups at neutral pH were used. Glutamic acid and aspartic acid carboxyl groups were deprotonated, and lysine ϵ-amino groups and arginine guanidine groups were protonated. Histidines were treated as neutral with the proton placed on the N^{δ} atom. N- and C-terminal residues were patched with neutral acetate and amide groups, respectively, as in the published structures. A shift function was used with a cutoff at 12 Å for the van der Waals interactions. The particle mesh Ewald method was used to treat the long-range electrostatics.[33] The real space contribution was

truncated at 12 Å. The number of grid points for the fast Fourier transformations was 81, 64, and 64 for the x, y, and z directions, respectively. Because the total charge of both the AB zipper and GCN4p1 was zero, no counterions were added. The simulations were performed in the isothermal, isobaric ensemble[34] by using the leapfrog integrator. The pressure and temperature were kept constant by using a Langevin piston of mass 600 amu and a Hoover thermostat with a thermal piston of mass 1000 kcal ps^2. All bonds involving hydrogen atoms were constrained by SHAKE[35], and an integration timestep of 2 fs was used in all simulations. Structures were collected every 0.5 ps for use in the pK_a calculations.

The coiled coils were immersed in a rectangular box of dimensions 70 × 50 × 50 Å containing preequilibrated TIP3P water.[36] The dimensions of the leucine zippers are approximately 42 Å long and 18 Å at the sides. Hence, the minimal distance from any protein atom to the edge of the box was 14 Å. Water molecules overlapping with protein heavy atoms (distance < 2.6 Å) were deleted. The total number of water molecules was 5210 for the AB zipper and 5178 for GCN4p1. All subsequent calculations were performed under periodic boundary conditions. The system was first relaxed by 50 steps of steepest descent minimization and then minimized for 200 steps by the adopted-basis Newton-Raphson (ABNR) algorithm with all water oxygen atoms and all protein heavy atoms harmonically constrained by a force constant of 1.0 kcal mol^{-1} $Å^{-2}$ and 2.0 kcal mol^{-1} $Å^{-2}$, respectively. After the minimization, the system was equilibrated for 30 ps under isothermal (298 K), constant volume conditions applying harmonic constraints on the protein heavy atoms with a force constant of 2.0 kcal mol^{-1}·$Å^{-2}$. The equilibrated system was heated to 298 K in 15 ps. The harmonic constraint force constant was progressively decreased from 2.0 to 0.0 kcal mol^{-} $Å^{-2}$ in the initial 10 ps, and all atoms were free during the last 5 ps of the heating process. The system was then subjected to a 20-ps equilibration at constant temperature and volume by using Gaussian distribution for the assignment of atomic velocities. The production simulations were performed for 0.3 and 0.7 ns (GCN4P1) or 1 ns (AB zipper) at 298 K. Details on the simulation of barnase have been published.[37]

Analysis of the MD Ensemble

Clustering of MD conformers

From each trajectory, coordinates were collected every 0.5 ps and were clustered on the basis of the combined backbone and charged side-chain heavy atoms root-mean-square deviation (RMSD) (excluding the first and last residues in each chain). This was expected to pool together structures not only close in global conformation but also displaying similar ionizable side-chain orientations. The cutoff value was 0.9 Å for the AB zipper and GCN4p1 and 1.0 Å for barnase. The clustering procedure followed published protocols.[38,39] The RMSD for the selected atoms was calculated for each pair of structures after optimal superposition. The number of neighbors was then computed for each structure by using the specified cutoff. The conformation with the highest number of neighbors was considered as the center of the first cluster. The second cluster was determined in the same way, excluding conformers that had been assigned to the first cluster. The procedure was repeated until each structure was assigned to a cluster.

The purpose of the clustering was to reduce the vast numbers of MD conformations to a fewer representative structures to be used in the pK_a calculations and, at the same time, to sample as many conformations as possible. Therefore, cutoff values of 0.8, 1.0, and 1.5 Å were also tested for the AB zipper. A 0.8 Å cutoff yielded too many clusters with only few members, so that the number of pK_a computations required was not significantly reduced. At the other extreme, a 1.5 Å cutoff produced too few clusters and, hence, less accurate description of the conformational variability. The 0.9–1 Å cutoff resulted in a reasonably small number of clusters, typically 3–6 per trajectory. Within a given trajectory, the major clusters (first, second, and third most populated clusters) grouped each 25–40% of all conformations, whereas the minor clusters represented each 6–8% of the population. Clusters populated to at least 5% of the total sampled conformations in each trajectory (e.g., 2000 for AB zipper) were used for the pK_a calculations, the rest being considered statistically insignificant. The 5% limit is somewhat arbitrary. However, the variation in the percent of discarded structures between 0 and 5% introduced statistically insignificant variation in the final results obtained for the ensemble (not shown).

Cluster-based averages

Coordinates at the center of the clusters as well as average coordinates within clusters were used for pK_a calculations (Table I). Different averages over the MD ensemble were built. According to the cluster-average protocols CS and CW, pK_a values were calculated by using coordinates representing the geometric mean of all conformations within a given cluster. The cluster-based trajectory average was then obtained by simple titration curve averaging (CS) or by population-weighted titration curve averaging (CW). Similarly, average pK_a values were calculated by taking the simple mean (CeS) or weighted mean (CeW) of the titration curves obtained for each ionizable group in the structure of the cluster center. The weighted sum of the titration curve for a charged group i,$\langle x_i \rangle$in CW and CeW can be calculated as $\langle x_i \rangle = \sum_{j=1}^{N} (w_j x_i)$ where x_i is the titration curve of group i in cluster j, w_j is the relative population of cluster j ($0 < w_j < 1$) and N is the number of clusters. In CS and CeS, w_j equals $1/N$. When an average over multiple trajectories is considered, it represents an arithmetic mean of $\langle x_i \rangle$ for all simulations.

Snapshot-based averages

From each trajectory, equally spaced conformations (snapshots) were extracted in 10-ps intervals and were used for pK_a calculations. The titration curves were then averaged over single trajectory or multiple trajectories

TABLE I. Comparison of the Various Sampling and Averaging Techniques Used for pK_a Calculations[†]

Abbreviation	Description	Sampling by[a]	Conformational averaging (over)[b]	Titration curve averaging (over)[c]	Trajectory/protein[d]
NMR	NMR conformers	NMR	No	Yes (20 or 25)	AB zipper, barnase
NAT	NMR average structure	Averaging	Yes (NMR)	No	AB zipper
XR	X-ray structure	X-ray coordinates	No	No	GCN4p1, barnase
CW	Cluster (weighted)	Clustering	Yes (clusters)	Yes (3–18)	All
CS	Cluster (simple)	Clustering	Yes (clusters)	Yes (3–18)	All
CeW	Cluster center (weighted)	Clustering	No	Yes (3–18)	All
CeS	Cluster center (simple)	Clustering	No	Yes (3–18)	All
SS	Snapshots	Snapshots	No	Yes (30–400)	All
TA	Trajectory average	Averaging	Yes (600–2000)	No	All
TA/4	Mean of 4 TAs	Averaging	Yes (2000)	Yes (4)	AB zipper
ATA	Total trajectory average	Averaging	Yes (8000)	No	AB zipper

[†]See text for a detailed description.
[a]The strategy used to sample conformations from MD trajectories and NMR ensembles.
[b]The number of conformations subjected to conformational averaging are given in parentheses. The number of conformations varies in each cluster.
[c]Numbers in parentheses indicate the total number of titration curves calculated before final averaging.
[d]Trajectories and/or proteins for which a particular sampling or averaging method is applied.

(30–400 snapshots of coordinates; protocol SS). This procedure is the same as the titration curve averaging method used previously by van Vljimen et al.[9]

Coordinate averages

First, the average structure representing all conformations in a particular trajectory was computed by geometrical averaging of the extracted coordinate sets (2000 for AB zipper, 600 or 1400 for GCN4p1, and 1400 for barnase). The pK_a values (titration curves) were then calculated on the resulting structure (protocol TA) as usual for single conformers. If the MD ensemble was obtained from multiple simulations, as for the AB zipper, for which four simulations were analyzed, the final MD-based average pK_a was the arithmetic mean of all individual TA estimates. This protocol is termed TA/4. Finally, according to the all-trajectory averaging procedure (ATA), one pK_a calculation was performed on the structure representing the mean of all 8000 conformations sampled in four independent MD simulations.

Other averaging techniques

For one of the simulations, which was started from an average of 50 NMR conformers (T4), pK_a values were calculated on the basis of the averaged electrostatic energy (including all contributions, i.e., ΔG_{solv}, ΔG_{bg}, and ΔG_{Wij}). No significant difference was found between pK_a values computed by such an averaging and pK_a values obtained by titration curve averaging techniques. This result is in line with the findings of van Vlijmen et al.[9] Therefore, all the other calculations followed the titration curve averaging method.

In a slight modification of the SS protocol, 20 conformers in a time window of 10 ps interval were averaged, and the resulting mean structure was used to calculate titration curves. The titration curves thus obtained were used to compute a trajectory average as in SS (protocol C_{10}). The results from C_{10} were very similar to those obtained from SS.

Analysis of the NMR Ensemble

The NMR structures representative for the solution structures of the AB zipper and barnase were analyzed in similar ways as described above for the MD ensemble (Table I). The protocol termed NMR was equivalent to snapshot averaging (SS). pK_a values listed in Table II as NAT were calculated on the average NMR structure in the same way as in protocol ATA. Necessarily, much fewer conformers (20–25) were subjected to averaging.

pK_a Calculations

All pK_a calculations were performed with the program MEAD.[10,40] The methodology has been described in the original publication.[10] Briefly, to estimate the change in the ionization behavior of a given group located on a protein with respect to the isolated group in solution, three types of energy contributions are calculated by numerical solution of the linearized form of the Poisson-Boltzmann equation.[41] The desolvation energy (Born energy, ΔG_{solv}) describes the difference in the interactions of the neutral or charged form of a group with the polarization that is induced when the group is located in the protein or in solution. The second electrostatic energy contribution ("background" energy, ΔG_{bg}) arises from interactions of the group with all other nontitratable charges in the system, such as partial charges distributed over all neutral protein groups and peptide dipoles. These two energy terms, which are both independent of the protonation state of all other sites and hence do not change with pH, are often combined with model compound pK_a values and referred to as the "intrinsic pK_a," pK_{intr}. pK_{intr} ostensively describes the change of electrostatic energy on the transfer of the considered group from the solution to its defined position in the protein, given all other ionizable groups are

TABLE II. Calculated Glutamic Acid pK_a Values From NMR and MD Ensembles of the AB Zipper Compared with Experiment[†]

Residue	Experiment[a]	NMR conformer average (NMR)	Average NMR structure (NAT)	Snapshot average (SS)	Weighted cluster average (CW)	Trajectory average structure (ATA)
E1	4.14	3.42	3.37	3.51	3.60	3.81
E6	4.82	5.24	5.87	5.11	5.02	4.77
E8	4.52	3.84	4.48	4.75	4.40	3.58
E13	4.37	4.33	4.94	4.64	4.38	4.08
E15	4.11	3.26	3.78	4.80	4.42	4.14
E20	4.41	4.11	3.52	4.50	4.51	2.08
E22	4.82	5.90	7.05	5.44	4.96	3.13
E27	4.65	4.00	3.78	4.56	4.26	4.38
E29	4.63	4.13	4.13	4.35	4.57	4.10
E1′	4.22	3.05	3.55	3.67	3.81	1.88
aad[b]		0.65	0.79	0.37	0.23	0.75
aad-T[c]		0.58	0.65	0.33	0.18	0.60
RMSD[d]		0.72	0.97	0.49	0.29	1.10
# conf.[e]		20	1	400	18	1

[†]Listed values are calculated with $\varepsilon_i = 4$. A detailed description of the abbreviations is given in the text and in Table I. The prime indicates residues in the basic strand of the two-stranded AB zipper. pK_a values for all titratable groups in the AB zipper calculated from all averaging protocols tested are listed in Table sI of Supplementary Material.
[a]Experimental pK_a values from Ref. 25.
[b]Average absolute deviation.
[c]Average absolute deviation excluding the N-terminal E1 and E1.
[d]Root-mean-square deviation.
[e]Number of titration curve calculations used to obtain the final pK_a.

neutral. The third energy contribution is the pH-dependent charge-charge interaction energy of each site with all other sites.

Once the pK_{intr} and the site-site interaction matrix are calculated, the protonation fraction f_i of each site i at a given pH is computed by Boltzmann weighted averaging over all possible protonation microstates of the protein.[10] The multiple-site titration problem is often intractable for a protein with large number of titratable sites. Various approximations have been devised to overcome this problem.[4,5,42,43] In this work, the reduced site approximation[4] was used for the AB zipper because the number of ionizable groups in this protein is relatively small (24).[4] The Monte Carlo (MC) procedure[42] has been used in the initial stage of the calculations for the AB zipper, and the results were the same as those obtained from the reduced site method. For GCN4p1 with 34 ionizable groups and barnase with 37 ionizable groups, the MC procedure was used.

The PARSE charges and atomic radii were used in all the calculations.[44] The ion exclusion radius was 2.0 Å and the probe radius was 1.4 Å. Ionic strength was set to the experimental values of 0.1, 0.15, and 0.05 M for the AB zipper, GCN4p1, and barnase, respectively. The solvent polarizability was modeled with a dielectric constant of 80. The internal protein dielectric constant has been shown to be a sensitive quantity in the continuum electrostatics model.[13–16] Different values for the protein interior dielectric constant, ranging from 3 to 20, have been tested. Most data reported for the AB zipper and GCN4p1 were obtained with $\epsilon_i = 4$. Following Dillet et al.,[22] the model pK_a (pK_{model}) were taken from Nozaki and Tanford,[45] namely 4.4 for Glu, 4.0 for Asp, 10.4 for Lys, 12.0 for Arg, 7.0 for His, and 9.6 for Tyr. All hydrogens, including carboxylate hydrogens, were built by using HBUILD.[32] The structures were then energy minimized by the CHARMM forcefield[30] by using the CHARMM22 all atom parameter set.[31] Before pK_a calculation, all structures were subjected to minimization, initially for 200 steps of steepest descent with the electrostatic function turned off, followed by 200 steps of conjugate gradient with the electrostatic term and using a dielectric constant of 4.0. A similar two-step minimization has been proposed as the best procedure in the pK_a calculations of lysozyme.[18] After minimization, the structures deviated by 0.4–0.6 Å all-atom RMSD from the starting structure. The final minimized structures were analyzed with PROCHECK to assess any deviations from the allowed geometry.[46] This was particularly important for structures representing averages of coordinates that could, in principle, contain unphysical geometries.

RESULTS AND DISCUSSION

Because an experimental pK_a represents a (fluctuating) system-average, whereas continuum pK_a calculations use structures with fixed atomic positions, computation of pK_a from an ensemble of structures requires suitable sampling techniques. The main ideas underlying the averaging and sampling protocols used in this study are summarized in Table I.

AB Leucine Zipper

The NMR ensemble

The disulfide-linked heterodimeric coiled coil AB (AB leucine zipper) consists of the acidic chain Ac-E-VAQLEKE-VAQAEAE-NYQLEQE-VAQLEHE-CG-NH2 and the basic chain Ac-E-VQALKKR-VQALKAR-NYAAKQK-VQALRHK-CG-NH2. All *e* and *g* heptad positions (underlined)

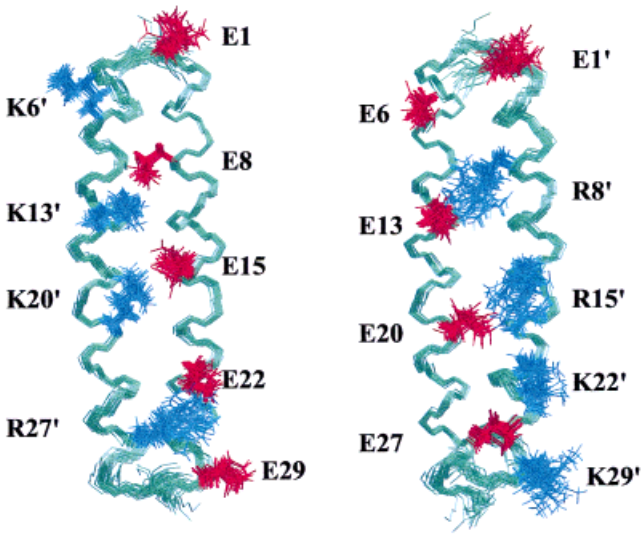

Fig. 1. Spatial orientation of charged *e* and *g* position residues in the ensemble of 25 selected AB zipper structures. Side-chain bonds of acidic and basic residues are displayed in red and blue, respectively, and backbone bonds are outlined in green. Front and back views of the zipper are shown.

are occupied by either E or K/R so that, by design, a maximum of six interhelical salt bridges can be formed. The solution structure of the AB leucine zipper at pH 5.6 has been reported, together with pK_a values for all Glu residues obtained from the pH dependence of proton chemical shifts between pH 1.5 and 7, where the AB zipper is fully folded.[25] The surface exposed, mostly charged residues, show a relatively defined alignment across the helix interface and partial H-bond formation between the functional groups of residues in *g* and *e′*(+1) of the adjacent helix is observed in the calculated structures. Still, there is a substantial variation in the orientation of the potential salt bridge partner residues toward each other (Fig. 1). The experimental pK_a values for all 10 Glu residues display no significant deviations from a model compound pK_a of 4.4, the largest upshift and downshift being +0.42 and −0.29 pK units, respectively (Table II).

For almost all Glu residues the pK_a values calculated on individual NMR structures fluctuate significantly across the NMR ensemble. pK_a values as high as 7.2 and as low as 1.8 are calculated for some groups. The conformational differences between the members of the NMR ensemble were taken into account by computing either the simple mean pK_a values over individual conformers or by calculating the pK_a on the average NMR structure. Because the backbone conformation of the AB zipper is well defined (RMSD 0.47 ± 0.13 Å), these two averaging protocols reflect mainly the local fluctuations of the electrostatic field in the vicinity of the charged side chains. The calculated pK_a values from the NMR ensemble are in reasonable agreement with the experiment (Table II). Large deviations are found only for few Glu residues. Inspection of the structure of individual conformers reveals no clear trend, which could explain why some conformations produce considerably larger errors than others. Because the mobility of surface exposed side chains is generally greater than that of the backbone, one would expect that the spread of pK_a values across the NMR ensemble mirrors the motion of neighboring and interacting side-chain charges. This is indeed the case for most of the designed ion pairs as well as for the N-terminal E1 and E1′. Typically, the conformer-related differences in the calculated pK_a values scatter within 1 pK unit around the experimental pK_a. This behavior is exemplified in Figure 2 for E8, the pK_a of which appears to be mainly influenced by the interaction with its ion-pairing partner, K13′. However, there are interesting exceptions. As shown in Figure 2, the interaction of E20 with the charged form of other titratable residues, mainly with its potential salt bridge partner R15′, changes little because the distance between $O_\epsilon1/O_\epsilon2$ of E20 and $N_\eta2$ of R15′ remains nearly constant throughout the 20 NMR structures. The desolvation contribution to the pK_a is almost constant as well (except for slight fluctuations in structures 16–19), suggesting a similar degree of charge burial in most structures. The fluctuations of the pK_a mainly arise from the interaction of the carboxyl group with background partial charges, particularly with the carbonyl oxygen of R15′. Similarly, the background interaction energy term, not the interaction with the potential salt-bridge partner R27′, affects the protonation equilibrium of E22 (Fig. 2). In this case, the pK_a varies in the same way as the distance between the closely positioned atoms $O_\epsilon1/O_\epsilon2$ of E22 and $O_\epsilon1$ of Q25. Not surprisingly, if single NMR conformers are used in the calculation, the largest absolute deviations from the experimental pK_a are calculated for E20 and E22. It is worth noting that the average calculated pK_a of E20, 4.11, is not far from the experimental value of 4.41 due to the low percentage of structures with extremely upshifted pK_a that balance the cumulative contribution from most conformers displaying slightly downshifted pK_a. On the other hand, E22 is poorly predicted because conformers with very high pK_a values dominate the ensemble. The pK_a values calculated by using the average structure (NAT) are less accurate compared with those from the ensemble. A larger error is obtained particularly for E20 and E22 (Table II).

Overall, titration curve averaging seems to be superior to conformational averaging when applied to the NMR ensemble and results in a fair prediction of the experimental pK_a, the average absolute deviation being 0.65 pK units (Table II). Six groups (E1, E1′, E8, E15, E22, and E27) are predicted with an error larger than 0.5 pK units (aad = 0.85). Three of these can potentially form salt bridges with lysine residues (E8, E15, and E27). E1 and E1′ are located at the N-termini of chain A and chain B.[25] The other four residues (E6, E13, E20, and E29) are predicted within an aad of 0.33 pK units. E6 and E29 are surface exposed in the NMR structures and thus might be very flexible, whereas E13 and E20 are located in the middle of the helix in a relatively fixed orientation and both have arginine as ion pair partners. Therefore, it appears that if the conformational variability of the AB zipper is accounted for by computing an

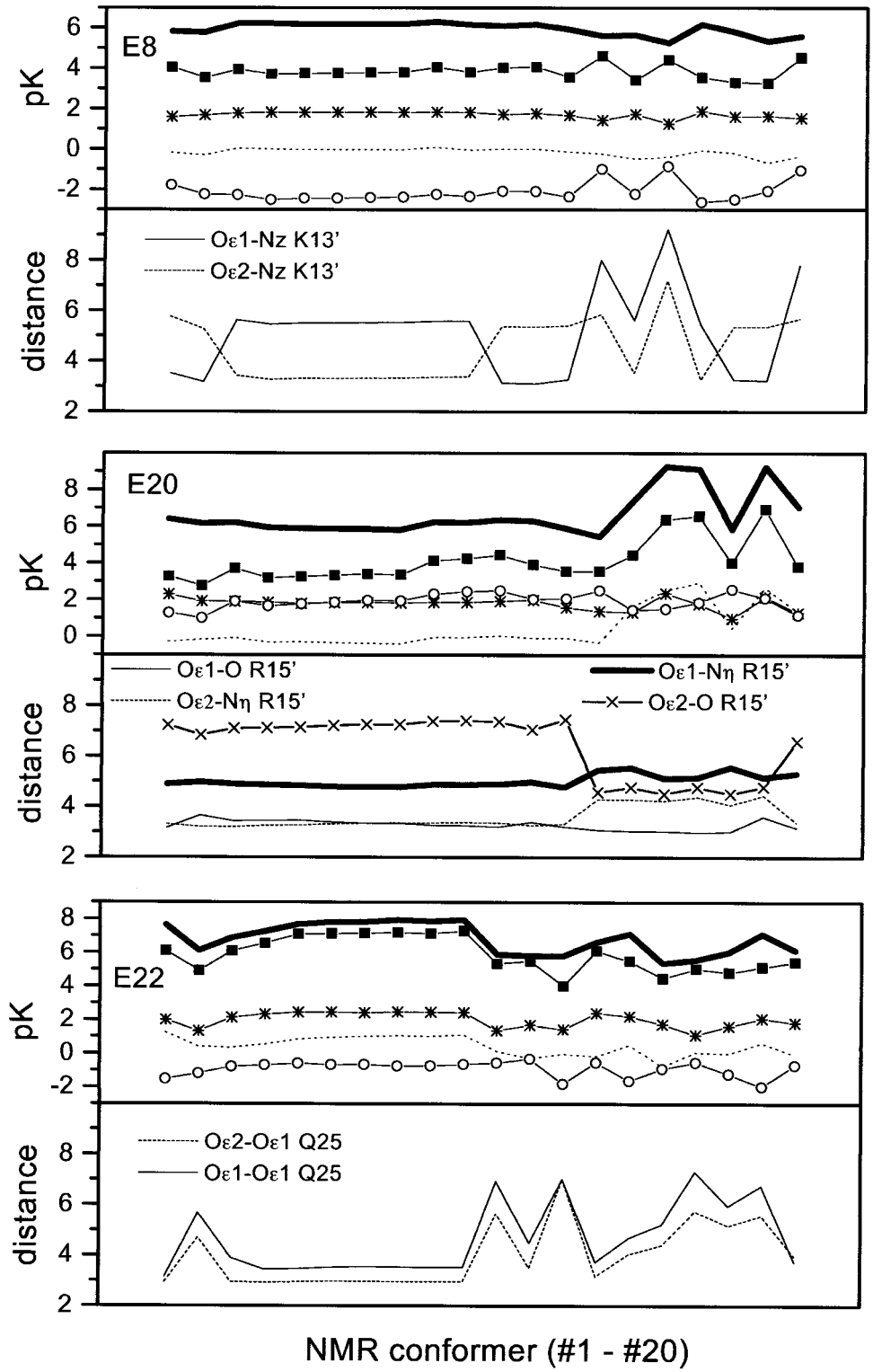

Fig. 2. Energy contributions to pK_a of selected Glu residues of the AB zipper calculated from the NMR ensemble. The upper part of the plots displays the final calculated pK_a (squares) across the NMR ensemble, together with the constituting energy terms. Open circles, energy of interaction with other titratable charges (Wij); asterisks, changes in solvation; dotted line, interaction energy with permanent partial charges (background term); thick continuous line, pK_{intr}. All energy terms were calculated with $\epsilon_i = 4$. The conformer-to-conformer pK_a variation correlates most strongly with the W_{ij} term ($R = 0.8$) in the case of E8. The background term dominates the calculated pK_a of E20 ($R = 0.91$), whereas differences in both desolvation ($R = 0.84$) and background interactions ($R = 0.77$) are the major source of pK_a variation for E22. In the lower part of the plots, the distance of Glu atoms $O_\epsilon1$ (light continuous lines) and $O_\epsilon2$ (dotted lines) to atoms of other protein groups is shown. For E8, the distance to Nz of K13′ is depicted. For E20, the distance of $O_\epsilon1$ (thick line) and $O_\epsilon2$ (crosses) to the carbonyl oxygen of R15′ is shown, in addition to the distances to $N_\eta2$ of the same residue. In the case of E22, the distance to $O_\epsilon1$ of Q25 is given.

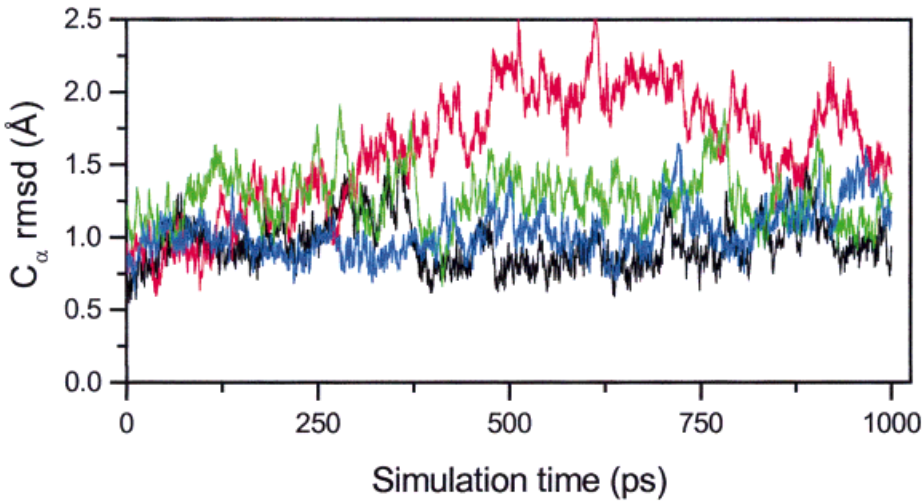

Fig. 3. Time evolution of C_α RMSD observed for trajectories T1 (in black), T2 (in red), T3 (in green), and T4 (in blue) of the AB zipper. RMSD is measured with reference to the corresponding NMR conformer used to initiate the simulation.

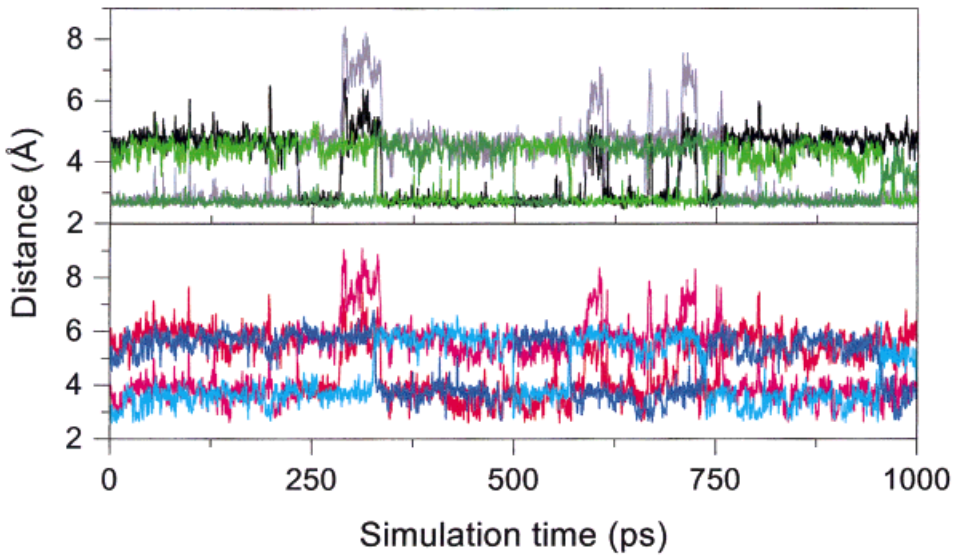

Fig. 4. Spatial interactions between groups involved in salt-bridge formation in the AB zipper observed by MD. The typical behavior of the E-R pairs during most of the simulation time is exemplified with the interatomic distances measured for the E13-R8′ pair in the course of simulations T3 and T4. Upper panel, $O_\epsilon1/N_\eta2$ (black) and $O_\epsilon2/N_\eta2$ (gray) in T3; $O_\epsilon1/N_\eta2$ (light green) and $O_\epsilon2/N_\eta2$ (dark green) in T4. Lower panel, $O_\epsilon1/N_\epsilon$ (red) and $O_\epsilon2$/Ne (magenta) in T3; $O_\epsilon1/N_\epsilon$ (blue) and $O_\epsilon2/N_\epsilon$ (cyan) in T4. Typically, three to four flips (rotation around the Cβ-Cγ bond) were observed in a single trajectory, consistent with the estimated timescale of surface-exposed side-chain rotation.[61]

average pK_a over the NMR ensemble, the accuracy of pK_a predictions is improved over single conformer calculations (including NAT) for Glu residues that are either distant in space from other positive charges and flexible or interacting with arginine residues in the middle of the helix. The discrepancy between calculation and experiment is larger for the N-terminal Glu and Glu involved in a potential ion pairing with lysines.

The MD ensemble

To further explore the effect of conformational variability on the precision of pK_a predictions, an ensemble of the AB zipper structures was generated by MD performed in explicit solvent. Three trajectories, T1, T2, and T3, each of 1-ns productive (sampling) duration, were started from three different NMR conformers that display a different pattern of salt bridges (E-K/R pairs present or absent). A fourth simulation, T4, was started from a conformation representing an average over 50 structures generated in the final NMR refinement protocol. Figure 3 shows the C-$_\alpha$ RMSD for the four simulations. After an increase in the initial 50 ps, the C_α RMSD calculated for simulations T1, T3, and T4 fluctuated around a mean value of 1.1 Å for the rest of the simulation time. In trajectory T2, C_α RMSD

increased further to about 2.2 Å but was reduced to 1.5 Å in the final 300 ps. All four trajectories converged during the last 200 ps and fluctuated around a mean C_α RMSD of 1.2 Å.

The overall conformational differences between the NMR ensemble and the MD-generated ensemble can be examined in terms of a pairwise RMSD.[47] The average of 20 NMR structures (NAT) was used as a reference for comparison. Individual starting structures differ from NAT by 0.43–0.64 Å (mean 0.56 Å) in C_α positions. The corresponding values of the individual trajectory averages are 0.83–1.25 Å (mean 1.05 Å). Within the MD ensemble, individual trajectories differ from each other by 0.46–1.1 Å. The global MD average structure is 0.91 Å away from NAT.

Several protocols were devised to account for the variability of conformation-related electrostatic properties of the AB zipper as observed in the MD simulation (see Table I). Table II compares experimental pK_a values to pK_a values calculated on the MD-generated ensemble and the NMR ensemble. Clearly, all averaging techniques applied to the MD-generated structures, with the only exception of ATA, significantly increase the accuracy of the predictions. In comparison with calculations with the NMR ensemble, both RMSD and aad are reduced at least twofold. The all-Glu prediction is further improved if one excludes the relatively poorer estimates for the N-terminal E1 and E1′ groups: the average error (aad) for these two sites is 0.5 pK units compared to an average aad of 0.2 pK units for the rest of the acidic groups. The more accurate pK_a calculation on the MD ensemble mainly arises from the improvement of the computed pK_a values for five sites, the pK_a values of which are relatively poorly predicted from the NMR ensemble (E1, E1′, E8, E15, and E27). In general, the combination of long and multiple MD simulations appears to generate an ensemble of structures representing the protonation equilibria of the Glu residues in the folded state of the AB zipper more realistically than the NMR ensemble.

In the continuum electrostatics model, the pK_a shift (from the model compound pK_a) of a titratable group located in the protein is calculated by summing up energetic contributions that are very sensitively dependent on the local geometry.[18,20] Local details necessarily reflect the global conformation and are not easily separable from it. This is especially true for small nonglobular domains lacking an extended hydrophobic core such as the leucine zipper. Therefore, we ask whether analysis of certain conformational features of the MD ensemble and their fluctuations in each individual simulation, as well as trajectory-to-trajectory fluctuations can explain the success of pK_a predictions computed from MD ensembles.

Four of the 10 Glu residues in the AB zipper, namely E1, E1′, E6, and E29, are very well solvated (relative solvent accessibility ≥ 0.7) and do not participate in interhelical salt bridges. Experimentally, E6 and E29 display a slightly upshifted pK_a, which is reproduced with high accuracy by the calculations by using both the NMR and the MD ensembles. Perhaps unexpectedly, the experiment indicates a pK_a downshift for the N-terminal E1 and E1′. In the NMR ensemble as well as in the course of the MD simulations, the carboxyl groups of both terminal Glu residues are closely approaching the amide nitrogen of Val 2 and Val 2′, respectively (mean distance 4.1 ± 1.0 Å). The persistence of such an interaction is in line with the observation that hydrogen bond acceptors and especially negatively charged side chains at the N-terminus of an α-helix might contribute to helix stabilization due to favorable interactions with the helix macrodipole.[48] Indeed, the computed pK_a values of these groups are low for almost all NMR conformers, averages thereof, and MD averages. The predicted pK_a shift is clearly dominated by the background energy term. The desolvation energy contribution is modest (0.2–0.6 pK units), whereas the interaction with other titratable charges is very small (approximately −0.2 pK units) because the carboxylate groups are solvent exposed and distant from other ionized side chains. From the MD ensemble, pK_a of E1 and E1′ are predicted with higher precision than from the NMR ensemble, possibly because of accumulation of thermally accessible conformers, in which the interactions with the backbone amide atoms are weakened. Most importantly, the pK_a values obtained for the MD averages, although lower than the experimental values, correctly reproduce the pK_a of E1 to be more downshifted than the pK_a of E1′.

The remaining six Glu residues participate in interhelical salt bridges. Although partial H-bond formation between Glu and Lys/Arg residues can be inferred from the solution structure (Fig. 1), the pH-dependent chemical shifts of Glu H^γ resonances and FDPB calculations indicate no pK_a downshift or even a slight pK_a upshift for most groups. All the six salt bridges were observed in the course of the combined simulation time of 4 ns ($O_\epsilon1/O_\epsilon2$–N_η/N_z distance < 3 Å). Their population varied from trajectory to trajectory and was independent of the presence or absence of the interaction in the starting structure. Some examples of the time evolution of salt bridging interactions are plotted in Figure 4.

Decomposition of the calculated pK_a values for salt-bridge participating Glu residues into various energetic contributions reveals a general trend that the charge-charge interaction energy is the major factor influencing the magnitude and the direction of the structure-to-structure pK_a variation (Figs. 5 and 6). As expected, the contribution of the positively charged salt-bridge counterpart dominates over the net influence of all other titratable groups. The predicted pK_a for different conformers display only low correlation to either the desolvation energy term or to the interaction energy with the partial charges and even lower correlation if these two terms are lumped together in the convenient manner to calculate the pK_{intr}. Nonetheless, the observed trajectory-to-trajectory variability of pK_a shifts is clearly governed by the background interaction term (Figs. 5 and 6). For most salt bridges the uncharged form of Glu is favored because of the often positive background contribution, irrespective of the percentage population of the salt bridge in a trajectory. It was shown above that the same effect was responsible for relatively high calculated pK_a values of E20 and E22 in the

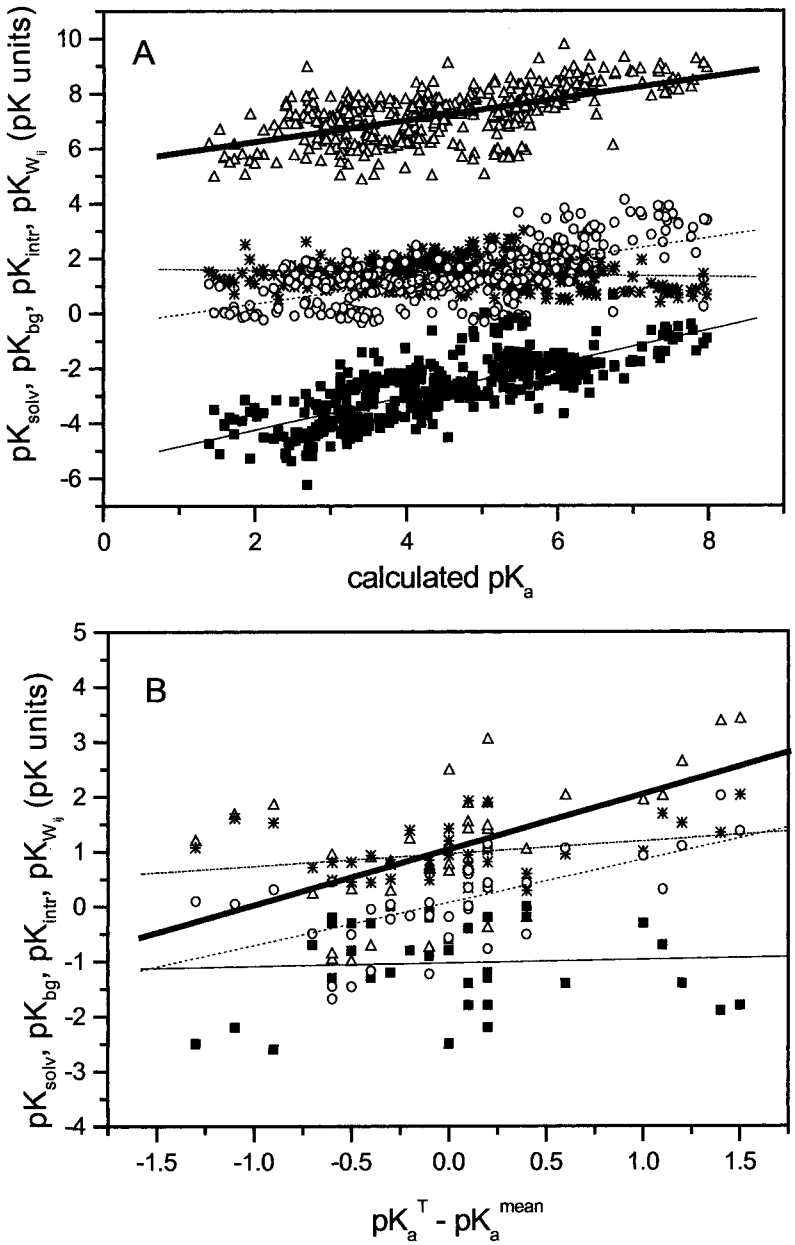

Fig. 5. Decomposition of calculated Glu pK_a values for the MD ensemble of the AB zipper into energetic contributions. All calculations were performed with $\epsilon_i = 4$. Squares, W_{ij} term; circles, background interaction; asterisks, solvation energy; triangles, pK_{intr}. The lines represent regression lines. Panel A combines the results of four MD simulations. For a total of 400 snapshots, calculated pK_a values best correlate with the magnitude of Wij ($R = 0.76$) followed by the background interaction energy ($R = 0.62$). There is no correlation with changes in solvation ($R = -0.12$). Panel B illustrates that the source of trajectory-to-trajectory pK_a variation is the variation of the background energy term ($R = 0.6$), whereas the correlation with the W_{ij} term is essentially zero.

NMR ensemble. As depicted in Figure 6, when the background term favors the charged form as in T4, or when it is small as in T1, the predicted pK_a values tend to be lower than the model pK_a of 4.4. In contrast, the high positive contribution of background interaction in T2 and T3 upshifts the pK_a values. Most of the experimental pK_a values, with the exception of E15, are higher than 4.4. Consequently, one could argue that conformations with a higher background term, as those dominating trajectories T2 and T3, are more frequent in solution. However, the predicted pK_a values from the latter two simulations are overestimated, even if the shifts are in the right direction. Therefore, only the global average over the MD-generated conformations, not each trajectory alone, yields pK_a values in close agreement with the experimental values. We believe this is an important result of the present work. It is tempting to speculate that conformations displaying positive or negative background energy contribution would be

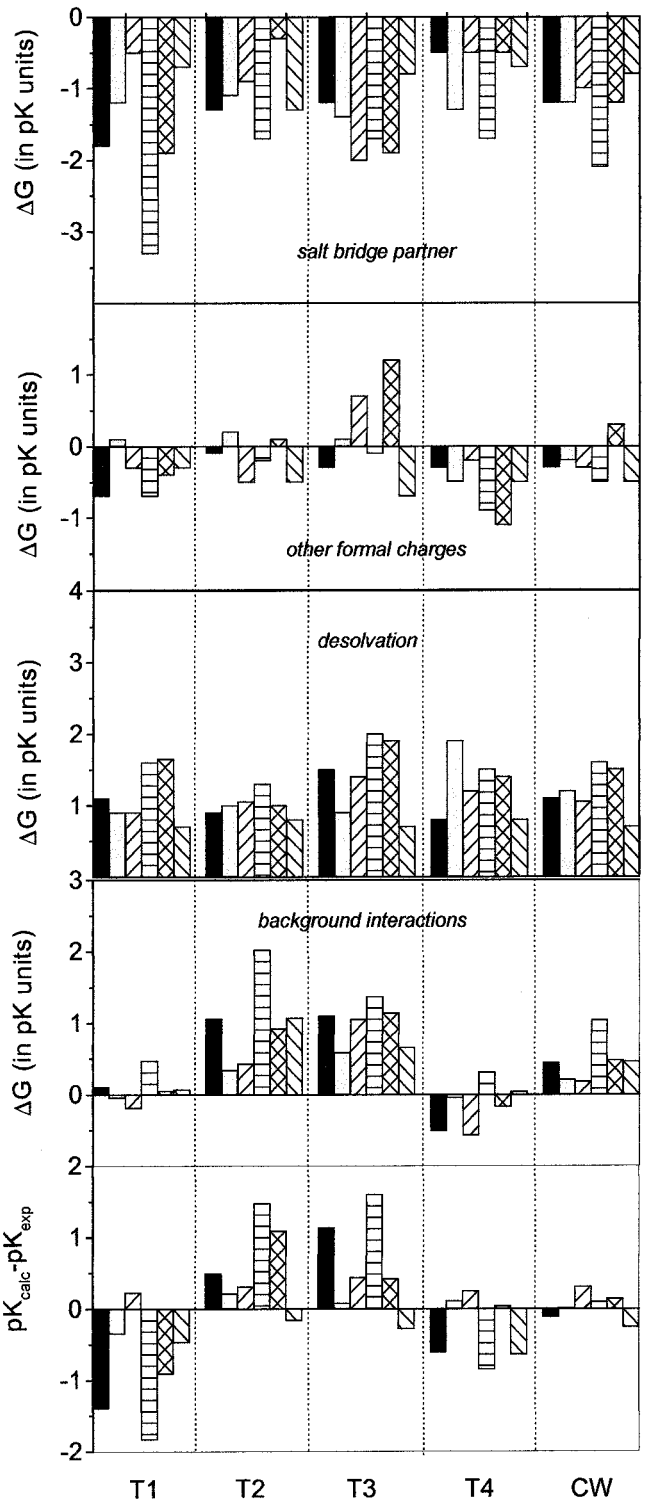

Fig. 6. The energetic contributions to the pK_a (upper four panels) and the error in pK_a calculation (lowest panel) for six Glu residues that participate in salt-bridge formation in the AB zipper. Vertical dotted lines demarcate trajectories T1 through T4. In the last segment on the right, results obtained from protocol CW are shown (see Table I and the text for details). The hatching pattern encodes for Glu residues E8, E13, E15, E20, E22, and E27 (left to right in each segment). Results for $\epsilon_i = 4$.

about equally populated in solution. Expressed differently, the four trajectories appear to sample nonoverlapping or partly overlapping regions of the electrostatic properties space.

The contribution of surface-exposed salt bridges to the stability of coiled coils is a controversial issue. Interhelical ion pairs have been described as either having no effect on stability or being slightly destabilizing[25,49,50] or being stabilizing.[51,52] The electrostatic free energy of salt bridges was shown to change sign at least once across 40 NMR conformers of the c-Myc-Max leucine zipper.[53] From the analysis of a large set of ion pairs, it was concluded that stabilizing or destabilizing effects are governed not only by the mutual geometrical orientation of the participating charged groups but are also very sensitively dependent on the exact distribution of full and partial charges in the protein matrix.[54] Therefore, conformational variability must be carefully assessed before classifying a salt bridge as being stabilizing or destabilizing based on continuum calculations.

Concluding this section on the AB zipper, we speculate that the correlation between the trajectory-to-trajectory pK_a values variation and the sign and magnitude of background interactions, in conjunction with an average of 1.1 Å C_α RMSD, might point to the importance of global conformational sampling in continuum electrostatics pK_a calculations. In the case of closely approaching charges on the surface of a protein, as in a typical salt bridge, local short-range side-chain movements predominantly result in fluctuation of the solvent-screened charge-charge interactions between neighboring full charges, which might not fully account for an experimentally observed pK_a shift. Extensive sampling is required to properly describe subtle fluctuations of the protein multipole around the "static" mean represented by an X-ray structure or an NMR conformer. Multiple nonconstrained MD simulations might be a useful alternative to fulfill this task.

GCN4 Leucine Zipper and Barnase

To test the general applicability of our averaging protocols and to further test the significance of MD based conformational sampling, we investigated two further protein domains, GCN4p1 and barnase. GCN4p1 represents the dimerization domain of the yeast transcription factor GCN4. The X-ray structure has been solved to 1.8 Å resolution, and there is abundant experimental and theoretical information concerning electrostatics-related properties of the molecule.[49,50,52,55] GCN4p1 is a coiled coil, the overall structure of which closely resembles that of the AB zipper.[25] Barnase is a typical globular protein with less uniform distribution of charges. Some acidic side chains are buried (relative solvent accessibility < 0.15), and three of them are situated in close spatial proximity in the active site.

GCN4p1

A best-fit superposition of the AB zippers' backbone heavy atoms of the four heptads with the corresponding atoms of GCN4p1 yields an RMSD of 0.85 Å. While the AB

TABLE III. Calculated pK_a Values From the Crystal Structure and MD Ensembles of GCN4p1[†]

		T1gcn4			T2gcn4		
Residue[a]	X-ray	Snapshots average (SS)	Weighted cluster average (CW)	Trajectory average (average structure) (TA)	Snapshots average (SS)	Weighted cluster average (CW)	Trajectory average (average structure (TA)
E20 (4.44)[b]	4.61	4.06	4.10	4.14	3.64	4.23	5.21
E20′	3.94	3.97	4.59	4.34	3.82	4.05	5.54
E22 (4.13)[b]	2.57	1.71	1.78	1.54	3.04	2.88	2.95
E22′	3.39	3.33	3.39	3.88	3.18	3.55	3.19

[†]Listed values are calculated with $\varepsilon_i = 4$. A detailed description of abbreviations used as column labels is given in the text and in Table I. pK_a values of all titratable groups calculated according to all averaging schemes tested are listed in Table sII in Supplementary Material.
[a]The prime indicates that the residue is located on the complementary chain of the homodimeric GCN4p1.
[b]Experimental pK_a for E20 and E22 is from Ref. 49 and represents the average pK_a obtained by NMR for E20 and E20′, and E22 and E22′, respectively.

zipper is a heterodimer, GCN4 is a homodimer, in which positively and negatively charged residues are located in the same peptide chain. Three salt bridges are evident in the X-ray structure of GCN4p1, E20-K15′, E22-K27′ and E22′-K27, corresponding to the E20-R15′, E22-R27′ and E27-K22′ interactions in the AB zipper. The pK_a values of E20 and E22 have been measured by ^{13}C-NMR.[49]

Two MD simulations, T1gcn4 and T2gcn4, which differ only in the initial assignment of the atomic velocities, were started from the X-ray structure. T1gcn4 was run for 700 ps and T2gcn4 for 300 ps, a total of 1 ns. The C_α RMSD fluctuated around a mean of nearly 1.2 Å, a value comparable with the one obtained in the simulations of the AB zipper. Following the same averaging protocols, pK_a values were calculated for the two trajectories, and the results are listed in Table III. It should be noted that, because this protein is a symmetrical homodimer, each experimental pK_a measured by NMR represents in fact a mean pK_a for two nondistinguishable sites. For the same reason, a single MD simulation provides alternative conformational sampling for each site. The experimental pK_a of E22 is 4.13, lower than the model compound pK_a of 4.4 and almost all of the Glu pK_a values in the AB zipper. Neither the X-ray structure nor any of the MD averages yield a correct prediction for E22, the average over the two trajectories and two identical E22 sites being 2.9. There are large differences between pK_a values obtained from T1gcn4 and T2gcn4 as well as between E22 and E22′. This indicates that larger sampling is required. From our simulations, the differences are caused not only by the varying strength of the interaction with the salt bridge partner K27 (and K27′) but also arise from interactions with two neighboring groups, R25 and H28. The presence of strong electrostatic interaction with these groups has been identified recently by others.[50]

The pK_a calculated for E20 from the X-ray structure and the two MD simulations are in reasonable agreement with the experimental value of 4.44. For most titratable groups, there is little variation in the pK_a between the trajectories themselves, and the correspondence between MD averages and the X-ray single-conformer calculation is fair. Still, for some groups (R1, Y17, R1′, E11′, R25′, and K27′), one or both simulations predict pK_a values significantly different from the values calculated on the GCN4p1 X-ray structure (>1 pK unit; see Table sII in Supplementary Material). The pK_a values of D7, E11, and H18 differ by >1 pK unit from pK_a values obtained for D7′, E11′, and H18′ when calculated from the crystal structure. In the MD ensemble, the predicted values for these pairs of symmetrical sites are much closer to each other. The results with GCN4p1 present yet another example that a given property space might not be effectively sampled by a single MD trajectory. The pK_a predictions listed in Table SII await experimental verification.

Barnase

Table IV lists experimental pK_a values obtained for barnase by NMR titration[56] together with pK_a values calculated by using several X-ray structures, 20 NMR conformers, and the MD ensemble from a recently published 700-ps simulation of the protein in explicit solvent.[37] Only charged groups with clearly defined experimental titration behavior can be analyzed with certainty. It should be noted that comparison with experiment is difficult for some residues listed in groups C and D by Oliveberg et al.[56] The authors mention the lack of well-defined titration baseline in some cases, resulting in overestimation of pK_a values as well as ionization concomitant with low pH unfolding.

Using $\epsilon_i = 4$, the deviation of the calculated pK_a values from the experimental values is generally not large with or without conformational sampling (Table IV). This is particularly so for groups listed in A and B by Oliveberg et al.[56] The aad is reduced from 0.8–1.0 for the single conformer calculations by using X-ray structures to 0.8 in the NMR ensemble and to about 0.6–0.7 in some MD averages. However, the major improvement comes mainly from E73. Excluding E73, most of the pK_a values are predicted reasonably well from the X-ray structures, and no significant improvement is achieved by any of the sampling techniques. For the entire set of residues, the deviation from experimental pK_a values is smaller for ϵ_i = 20, with only marginally better prediction being obtained for the MD averages. A high dielectric constant has been frequently assigned to the protein domain in continuum pK_a calculations and is thought to implicitly approximate the effect of structural flexibility on the ionization properties of titratable groups.[6] Are the results obtained for

TABLE IV. Experimental and Calculated pK_a Values for Asp, Glu, and His Residues of Barnase Obtained From X-ray Structures, 20 NMR Conformers, and MD Ensemble From a 700-ps Trajectory[†]

Residue	Experiment[a]	ε_i	X-ray (XR)[b]	NMR conformer average (NMR)	Snapshot average (SS)	Cluster center average (CeS)	Trajectory average structure (ATA)
D8	3.30	4	2.50–3.20	2.83	1.90	4.10	2.10
		20	2.90–3.30	3.23	2.50	3.30	2.50
D12	3.80	4	3.30–4.40	4.11	3.50	3.40	3.30
		20	3.20–3.70	3.75	3.30	3.40	3.30
D22	3.30	4	2.80–3.20	3.35	3.60	3.00	2.80
		20	3.00–4.40	3.63	3.10	2.90	2.90
E29	3.75	4	3.90–4.30	3.02	3.30	3.10	3.60
		20	3.80–4.40	3.84	3.60	3.60	3.70
D44	3.35	4	3.30–3.40	3.29	3.50	3.40	3.40
		20	3.50–3.60	3.63	3.60	3.50	3.60
D54	2.20	4	(−0.2–1.4)	1.63	1.50	1.50	4.20
		20	1.20–2.00	2.60	1.90	2.20	2.60
E60	3.00	4	1.80–3.40	3.20	3.40	3.40	2.70
		20	2.90–3.40	3.45	3.40	3.40	3.10
E73	2.20	4	4.70–7.10	5.79	1.20	3.30	0.50
		20	2.90–3.50	4.50	2.20	2.20	1.40
D75	2.80[c]	4	<-1.0	2.38	<-1.0	<-1.0	<-1.0
		20	0.30–1.00	2.43	0.40	0.30	0.60
D86	4.20	4	4.20–4.90	4.97	3.40	4.10	4.60
		20	3.10–3.30	3.83	3.20	3.20	3.50
D93	0.70[c]	4	(−0.9–3.20)	3.25	1.90	0.70	1.40
		20	1.70–2.90	3.20	2.10	2.20	2.30
D101	2.00	4	1.00–3.00	3.13	2.60	2.60	2.40
		20	2.30–3.00	3.43	2.90	2.90	2.90
H18	7.73	4	7.50–8.00	7.13	7.60	7.50	7.50
		20	7.40–7.50	7.20	7.40	7.40	7.40
H102	6.30	4	5.90–6.30	5.86	6.50	6.60	6.20
		20	6.60–6.80	6.87	6.70	6.70	6.60
aad[d]		4	0.86	0.77	0.73	0.60	0.79
		20	0.64	0.69	0.65	0.60	0.67
aad-2[e]		4	0.68	0.65	0.52	0.52	0.63
		20	0.45	0.57	0.44	0.35	0.46

[†]Listed residues were selected for analysis as explained in the text. A detailed description of the abbreviations used as column labels is given in the text and in Table I. Calculated pKa values from all averaging protocols tested are listed in Table sIII in Supplementary Material.

[a]From Ref. 56, except for H18 and H102, taken from Refs. 59 and 60, respectively. Where two or more experimental pK_a values are available from β/γ proton chemical shifts, we used the one with the lowest error.

[b]pK_a values calculated from five different X-ray structures are given as ranges, lowest to highest, whereas the aad is listed as the mean aad over all structures.

[c]D75 is not fully protonated below pH 2, so the actual pK_a might be much lower. Listed pK_a for D93 is the best estimate from mutational analysis at $\mu = 50$ mM.[56]

[d]Average absolute deviation. When the calculated pK_a of D75 was below zero, $pK_a = 0$ was used to calculate aad.

[e]Average absolute deviation excluding D75 and D93.

barnase supporting this idea? Overall, the answer is yes. The all-site average absolute deviation is significantly reduced for the X-ray structures on increasing ϵ_i from 4 to 20, becoming approximately equal to the aad of the MD ensembles at $\epsilon = 4$. Moreover, the accuracy of prediction is improved to a lesser extent for the NMR ensemble and the MD ensemble. Because the NMR and MD averages are less sensitive to ϵ_i, it appears that, in the case of barnase, explicit conformational sampling is a useful alternative to account for the effect of structural fluctuations on the ionization properties of charged residues. Thus, the need of introducing a further level of approximation in the continuum electrostatics methodology (i.e., the use of high dielectric constant) can be avoided. However, it is not possible to draw a general conclusion from the results obtained with barnase. This is evident if predictions for individual sites are considered. Most of the pK_a values are reasonably close to the experimental values when calculated from the X-ray structure at $\epsilon_i = 4$. For such groups, there is not much gain in accuracy from either using $\epsilon_i = 20$ or from the sampling. This is particularly true for surface-exposed residues such as D22, D44, E29, and E60 (groups A and B in Ref. 56). In some crystal structures, error of pK_a might even be larger at high ϵ_i (e.g., D22 and D86). A similar effect has been observed for lysozyme by others.[9] It is not surprising that buried groups and groups displaying coupled titration behavior (groups C and D in Oliveberg et al.[56]) are much more sensitive to the choice of dielectric constant and conformational variability. Table IV reveals some instructive extremes. pK_a of E73, perhaps

the most critical residue in the set, is correctly predicted only from the MD ensemble. Notably, however, a combination of titration curve averaging and higher dielectric constant was required to reproduce the experimental value. Obviously, sampling that is accumulated in a single trajectory, even a relatively long one, misses some important aspects in the fluctuation of the local electrostatic field around this particular site. The error for D93, on the other hand, is lowest at $\epsilon_i = 4$ from either two X-ray structures (other structures producing significant errors) or from MD averages. All averaging protocols with $\epsilon_i = 20$ result in poor prediction. It is likely, therefore, that for this site, titration curve averaging in conjunction with high-dielectric calculations accounts implicitly "twice" for contributions arising from the variation of the charge distribution in the ensemble present in solution. It is worth noting that the NMR ensemble average is superior to all other estimates in some cases (e.g., D75) for which all other calculations fail to reproduce the experimental pK_a. Taken together, the results for barnase imply the following. (i) Explicit conformational sampling, assessed from NMR or MD, improves the accuracy of the pK_a calculation in all-site-prediction terms using low values of ϵ_i. (ii) The ensemble-based calculations are less sensitive to the choice of ϵ_i and consistently reproduce highly shifted pK_a values at a low ϵ_i better than what might be achieved in a single-conformer calculation at high ϵ_i. (iii) For only a few residues, conformational sampling is critically required; most of titratable groups are predicted reasonably well by standard protocols. (iv) Single MD trajectories might sample only a small portion of the property space and, therefore, might inadequately describe the ionization properties of some groups. (v) Averaging techniques applied on NMR ensembles or MD ensembles might grasp different aspects of the conformational flexibility as a factor influencing continuum electrostatics calculations.

Comparison of Averaging Techniques

One goal of this work is to investigate the importance of MD-based extensive conformational sampling as an approach to increase the accuracy of pK_a predictions. A long MD simulation produces, in principle, a large number of conformations. Obviously, titration curve calculation using thousands of structures is not affordable in practical terms. On the other hand, conformers that are very close on the time coordinate are likely to be located virtually in the same point of the conformational space. Hence, they are likely to display negligible variation in electrostatic properties, so that calculated pK_a differences would eventually fall below the threshold of precision inherent in the continuum method. Therefore, several different protocols were devised in an attempt to replace the MD ensemble by a representative set of structures, without diminishing the conformational diversity of the ensemble. In the following we discuss the different averaging protocols in comparison with single conformer calculations and calculations on the NMR ensemble.

Titration curve averaging versus coordinate averaging

MD trajectories simulate the thermal motions of a solvated molecule around of (or away from) its mean conformation obtained by X-ray diffraction or NMR spectroscopy. Therefore, it is important to compare pK_a values calculated by making use of conformational averaging or titration curve averaging with pK_a values calculated from the starting structure. We take as an example trajectory T4, which was started from an average of 50 NMR conformers of the AB zipper and, hence, is not biased toward any particular NMR conformer. Table VA summarizes the analysis in terms of aad and RMSD. The improvement in aad achieved by all averaging techniques but TA ranges from 20% (SS) to 47% (CeS). Although the average structure (TA) slightly improves RMSD, aad remains the same as for the starting structure. In contrast, titration curve averaging markedly increases the precision. The same is true for trajectories T1 and T3 as well, which were started from individual NMR conformers (data not shown). Moreover, titration curve averaging is more accurate than conformational averaging also for barnase, whereas comparison is not possible with only two available experimental pK_a for GCN4p1. This is in contrast with the work of Van Vljimen et al.,[9] who have reported that the average structure from a 100-ps MD simulation of lysozyme and BPTI yields similar results as titration curve averaging. However, their simulations were short (100 ps) with respect to the ones of the present study, and the sampling might have been limited. The benefit of taking an ensemble average at the level of titration curves can be traced in the exceptional trajectory T2, where protocol TA yields an improvement of 45% in both aad and RMSD (data not shown). This particular trajectory had the highest C_α and all-atom RMSD in the set (Fig. 3) and displayed significantly higher fluctuations. Consequently, T2 might have explored a broader region of the property space, and the average structure appears to contain information ("memory") about more accessible states in that space, meaning that the improvement arises from the extent of sampling, not from the quality of the average conformation. The fact that in most of the MD simulations analyzed here titration curve averages perform better than the average structure indicates the strength of explicitly accounting for the conformational variability in pK_a predictions. The very same conclusion follows from the analysis of the NMR ensembles: titration curve average over NMR conformers yields better accuracy than the mean NMR structure (compare protocols NMR and NAT in Table II). In general, a single average conformation, even from a long trajectory, might not adequately represent the electrostatic properties of a protein and its fluctuations at room temperature.

Cluster-based pK_a calculations

In the light of the foregoing, the question arises whether the conformations in the MD ensemble are statistically fluctuating around a time-averaged constellation of the molecular multipole or whether some type of conforma-

TABLE V. Improvement in the Accuracy of pK_a Predictions From MD Ensembles of the AB Zipper[†]

A. Single MD simulation versus starting NMR conformer

	N4[a]	SS	CW	CS	CeW	CeS	TA	C10[b]	$E_{ave,CW}$[c]	$E_{ave,SS}$[d]
aad	0.76	0.61	0.44	0.50	0.54	0.40	0.76	0.60	0.46	0.63
%improvement	—	20	42	34	29	47	0	21	40	17
RMSD	0.97	0.73	0.56	0.64	0.64	0.56	0.83	0.71	0.58	0.75
%improvement	—	25	42	34	34	42	14	27	40	21

Results obtained with the 1-ns trajectory T4 in terms of aad and RMSD.

B. "Short" versus "long" MD simulations

	T1		T2		T3		T4		T_{tot}[f]	Mean[g]
Time (ps)	250	1000	250	1000	250	1000	250	1000	250 × 4	1000/4
aad	0.64	0.46	0.89	0.62	0.65	0.53	0.57	0.61	0.58	0.56
RMSD	0.87	0.56	1.09	0.75	1.10	0.70	0.70	0.73	0.71	0.69
aad-2[h]	0.33	0.39	0.53	0.36	0.33	0.38	0.48	0.48	0.38	0.40
Calculated pK_a										
E20 (4.41)[i]	5.79	4.11	6.94	6.14	4.29	4.44	3.49	3.31	5.13	4.50
E22 (4.82)[i]	5.53	5.20	5.92	5.66	8.01	6.30	4.82	4.57	6.07	5.43

pK_a values were calculated by using protocol SS.

C. Error from multiple MD trajectories versus the mean error from the corresponding starting structures

	N/4[j]	SS	CW	CS	CeW	CeS	ATA
aad	0.70	0.37	0.23	0.28	0.27	0.33	0.27
% improvement		47	67	60	53	61	61
RMSD	0.81	0.49	0.29	0.41	0.32	0.36	0.33
% improvement		40	64	50	56	60	59

[†]A detailed description of the abbreviations SS through ATA used as column labels is given in the text and in Table I.
[a]Calculation on the NMR structure used to start trajectory T4.
[b]Protocol C_{10} uses conformational averages over 10-ps intervals instead of snapshots for titration curve calculation.
[c]$E_{ave,CW}$ involves averages on the level of the electrostatic energy within clusters.
[d]$E_{ave,SS}$ involves electrostatic energy averages for snapshots.
[f]aad and RMSD calculated for the MD ensemble generated by combination of four trajectories, each of 250-ps duration.
[g]Average error over the 1-ns simulations.
[h]aad excluding E20 and E22.
[i]Experimental pK_a values in parenthesis.
[j]Denotes the mean aad or RMSD of the pK_a values calculated from structures N1–N4.

tions dominate over others with respect to their electrostatic properties. To gain information on this problem, we devised several clustering schemes. Because both local and global structural variability are explored by MD, clusters were built by taking into account fluctuations of side chains combined with fluctuations of the backbone. The clustering protocol resulted in three to six clusters per trajectory. At the level of individual trajectories, cluster-based averages perform better or equal to snapshots averaging. For the entire MD ensemble generated by multiple simulations, however, clustering is clearly superior to all other protocols. We do not notice a significant advantage of population-weighted averages over the arithmetic mean (see Tables sI–sIII in Supplementary Material). The likely reason is that for the proteins studied here the cluster-to-cluster pK_a variation is relatively small, perhaps because of the flexibility of most of the considered ionizable groups. We believe, in certain systems and/or for certain residues subjected to more stringent conformational constraints resulting in side-chain motions on a long timescale, statistical weighting of pK_a might be a useful way to overcome the limitations of single conformer calculations.

Cluster-based pK_a calculations could involve pure titration curve averaging (as in CeW and CeS) or a combination of titration curve averaging and coordinate averaging (as in CW and CS). The two methods yield comparable results suggesting that the average cluster conformation is very similar to the conformation of the cluster center. The main advantage of the clustering technique is the dramatic reduction of the number of pK_a calculations necessary to enfold the fluctuations of charges in the system while preserving or even increasing the accuracy of pK_a predictions.

Short versus long MD simulations

Because MD simulation of proteins in explicit solvent is computationally demanding, we examined how the length of the MD trajectory influences the pK_a predictions. One hundred snapshots, equally spaced over the initial 250 ps of simulation time, were used to calculate average titration curves for the AB zipper. The predicted pK_a values are closer to the experimental values for all 10 Glu residues of the AB zipper compared with the starting NMR structures. Long simulations improve the accuracy by 28%, 30%, and 19% in T1, T2, and T3, respectively, whereas

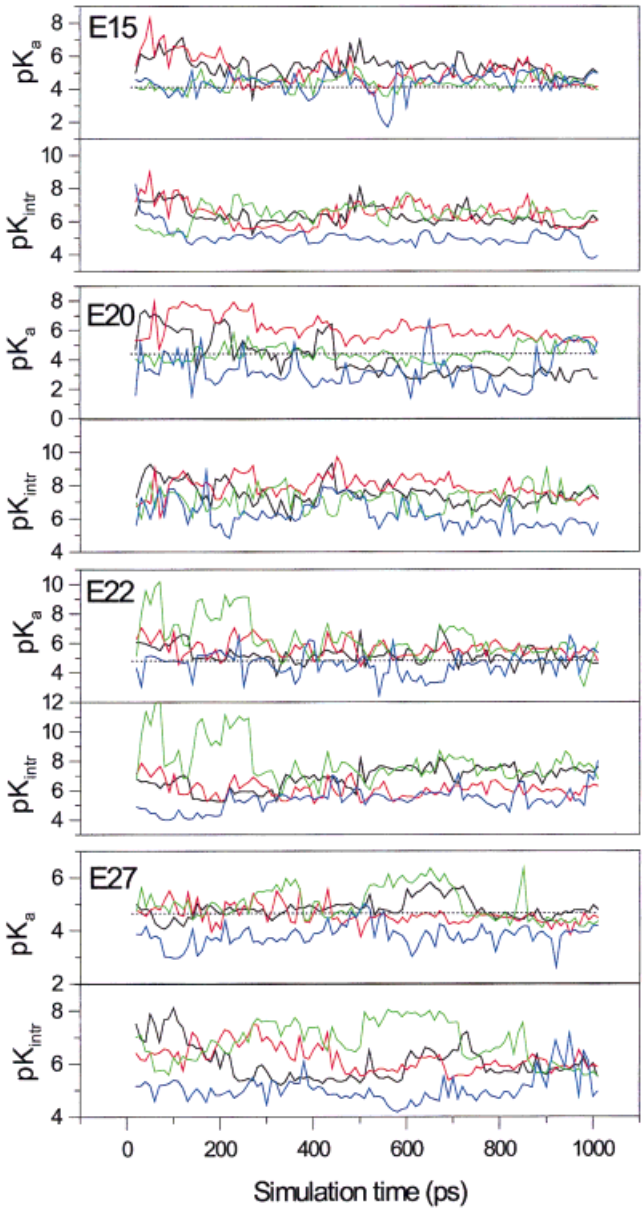

Fig. 7. Time evolution of calculated pK_a and pK_{intr} for Glu residues involved in potential salt-bridge formation in the AB zipper. Results obtained for 100 equally spaced snapshots from trajectories T1 (black), T2 (red), T3 (green), and T4 (blue) are shown. The dotted line in the upper panels is the experimental pK_a.

there is no further improvement in T4 (Table VB). Fluctuations of calculated pK_a values in the initial phase (approximately 300 ps) are mostly larger and tend to progressively attenuate toward the end of the simulation (Fig. 7). However, this is not the reason why a longer simulation reduces the error. For 8 of 10 Glu residues of the AB zipper, the mean pK_a is very similar in "short" and "long" simulations, except for the larger variation around the mean pK_a in the first 250 ps. The gain in accuracy of the overall prediction in the remaining 750 ps comes from only two residues, E20 and E22, which are located in the middle of the coiled coil and are surrounded by many charges. It appears, therefore, that for most of the surface-exposed charged groups, whether involved in salt-bridge interactions or not, short MD runs provide sufficient sampling for the pK_a calculations. Longer simulation times are required to reduce the error for only a few groups.

Single versus multiple simulations

One of the principal findings of the present work is that the combination of multiple MD simulations produces the most accurate pK_a values. In terms of overall aad and RMSD, single 1-ns trajectories of the AB zipper are superior to single conformer calculations on the starting NMR structure, yet produce the same or only slightly better accuracy as the NMR ensemble average. As a qualitative description of the outcome of single-trajectory pK_a predictions, we calculated for each Glu residue the difference $|pK_a^T - pK_a^{exp}| - |pK_a^N - pK_a^{exp}|$, which is the absolute deviation of the pK_a of the starting NMR conformer (pK_a^N) from the experimental pK_a (pK_a^{exp}) compared with the corresponding deviation of the trajectory-derived pK_a (pK_a^T) from the experimental value. We use an arbitrary cutoff of ± 0.2 to classify the pK_a values as "improved" (< -0.2), "unchanged" (-0.2 to 0.2) or "worsened" (>0.2) (Table VI). By using this cutoff of ± 0.2 pK units (which is about the largest possible error of the experimental pK_a values) 40% of the Glu pK_a values are improved on average in a given trajectory, 17% show a decreased accuracy, and the rest do not change. If a more relaxed criterion of 0.5 pK units is applied, which is usually considered as the margin of a reasonable pK_a prediction within the continuum method, about 50% of the groups fall in the "unchanged" category and only 12% are predicted more accurately from the starting structure than from the single trajectory. Even if few pK_a values are more accurately computed in all trajectories, the prediction for some other groups is closer or farther to the experiment than the NMR starting conformer, depending on the particular simulation considered. In contrast, the use of multiple trajectories results in a remarkable increase in accuracy, ranging from 47 to 67% aad and from 40 to 60% RMSD compared with the mean pK_a from the starting structures (Table VC).

Because single trajectories do not improve the accuracy of the all-site prediction over the NMR ensemble to a significant degree, the success of the ensemble generated by combination of multiple trajectories is quite remarkable. Importantly, even with large errors for E20 and E22, combining four independent short (250 ps) simulations started from different NMR conformers improves the accuracy of pK_a prediction as much as each long (1 ns) trajectory. The same is true for two trajectories started from the same initial structure of GCN4p1. For example, the pK_a values for E20 and E22 obtained from the initial 250 ps of trajectories T1gcn4 and T2gcn4 is 3.82 and 2.65, respectively, which are, on average, better than pK_a values calculated from the 700-ps simulation. In most cases, especially for residues located in the middle of the coiled coil which remain within H-bonding distance throughout the simulation, the final calculated pK_a from the MD

TABLE VI. Comparison of Glu pK_a Values of the AB Zipper Calculated From Individual NMR Structures or From 1-ns MD[†]

Residue	Experiment[a]	N1	T1	N2	T2	N3	T3	N4	T4
E1	4.14	3.67	3.89	3.04	3.79	3.69	3.50	3.71	2.87
E6	4.82	4.73	5.50	4.90	5.39	5.02	4.96	4.96	4.59
E8	4.52	2.83	4.86	3.80	5.13	4.04	5.48	2.28	3.55
E13	4.37	4.65	4.87	4.61	4.49	3.99	4.70	4.78	4.49
E15	4.11	5.95	5.40	3.68	5.05	3.91	4.37	4.27	4.39
E20	4.41	3.49	4.11	4.24	6.14	3.28	4.44	3.33	3.31
E22	4.82	4.98	5.20	7.28	5.68	6.12	6.30	6.22	4.57
E27	4.65	3.65	4.81	3.74	4.59	3.82	5.02	4.08	3.82
E29	4.63	4.62	4.40	4.09	3.97	4.16	4.71	4.23	4.33
E1′	4.22	3.06	3.74	2.93	4.41	2.86	3.20	3.33	3.32
Improved[b]			5		4		4		3
Worsened[b]			1		3		1		2
Unchanged[b]			4		3		5		5

[†]N1–N4 denote the NMR conformer used as starting structure for trajectories T1–T4, respectively. The numbers listed in columns T1–T4 were calculated according to the SS protocol (see Table I for details).
[a]From Ref. 25.
[b]Residues were classified as "improved," "unchanged," or "worsened" on the basis of the magnitude and sign of the difference $|pK_a^T - pK_a^{exp}| - |pK_a^N - pK_a^{exp}|$, which is the difference in the absolute deviations from the experiment. Values < -0.2 are considered an improvement, whereas values > 0.2 indicate a worse prediction from the trajectory ensemble. The $|0.2|$ margin is arbitrary (see text for details).

ensemble represents the mean of single-trajectory estimates, which scatter significantly around the experimental pK_a. Generally, there is no obvious dependence of the final pK_a calculated from the single-trajectory ensemble on the pK_a of the starting structure (the basis structure), as follows from the pairwise comparison of the data in Table VI. At the level of individual pK_a valuess, the trajectory-to-trajectory variation thus appears to arise mainly from the fact that the simulations explore different or only partly overlapping regions in the conformational space. It cannot be completely ruled out, however, that certain features of the starting conformers are preserved across the generated ensemble.

The pK_a and pK_{int} for some of the Glu residues involved in potential salt bridges are plotted in Figure 7 as a function of the simulation time. The pK_a values and pK_{int} of E15, E20, E22, and E27 fluctuate more during the initial 200 ps. Afterward, fluctuations decrease and the pK_a values converge. The most important information to be obtained from the plots in Figure 7 is that each trajectory, taken separately, can introduce one-directional error. For the same glutamates, some trajectories predict a downshift of the pK_a, some predict an upshift, and some predict no shift at all. Therefore, multiple trajectories are clearly required for proper sampling of most ionizable groups. In fact, worsening of pK_a accuracy has been reported for D66 and E7 of hen egg white lysozyme (HEWL) at $\epsilon_i = 4$ in an ensemble generated by a 100-ps simulation, compared with the X-ray structure.[9] Similar behavior is observed for some groups in the single trajectory-based ensembles of barnase and GCN4p1. As outlined by van Viljmen et al.,[9] the interaction between D66 and Y53 of HEWL, which appears responsible for the inaccurate pK_a prediction for D66, increases during simulation, thus increasing the error in the pK_a. The authors suggested that more sampling could resolve the problem. Although extended simulation times might provide a more detailed description in certain "problematic" regions, our results indicate that multiple short MD simulations could be possibly superior to single long trajectories.

Dielectric constant and conformational flexibility

In the framework of the continuum electrostatics method, the protein internal dielectric constant models the polarizability within the protein by a fixed uniform value. The use of high dielectric constants, which are physically less realistic, has been shown to give generally better pK_a predictions.[6] However, the use of a low dielectric constant has also been required to reproduce some highly shifted pK_a.[57] The parameterization of the protein dielectric within the continuum method has been studied in detail.[14,15,57] Any uniform value of ϵ_i represents an approximation accounting for the dipolar, atomic, and electronic polarizability (fluctuations). Hence, it is expected that the sensitivity of the calculated pK_a values on the exact value of the protein internal dielectric constant will decrease when conformational flexibility is considered explicitly. Indeed, pK_a values calculated from individual trajectories and from the entire MD ensemble of the AB zipper display a small variation between $\epsilon_i = 3$ and $\epsilon_i = 20$ and the accuracy of the all-Glu prediction is maximal at $\epsilon_i = 4-6$. In contrast, the accuracy of the pK_a values from single conformer calculations continuously increases from $\epsilon_i = 4$ to $\epsilon_i = 20$. As an example, Figure 8 displays the results for the trajectory T4 in comparison with its starting structure. The improvement in accuracy of pK_a calculations from MD-derived ensembles has been demonstrated to be highest at $\epsilon_i = 4$ for BPTI and HEWL.[9] Taken together with aforementioned results obtained for barnase at low and high dielectric constants, these observations indicate that computation of titration curve

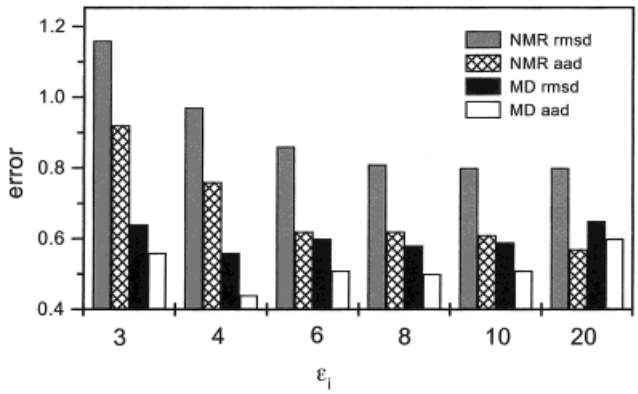

Fig. 8. Variation of the accuracy of pK_a predictions from NMR and MD ensembles of the AB zipper with ϵ_i. The error pK_{calc}–pK_{exp} is given in terms of RMSD and aad. Results obtained with trajectory T4 are shown. The same trend was observed with all other simulations, except that the minimal error was sometimes calculated for $\epsilon_i = 6$.

averages from MD trajectories at relatively low ϵ_i provides a physically more appropriate approach for the description of the ionization behavior of titratable groups in proteins.

NMR ensembles versus MD ensembles

The need of accounting for the conformational variability of a native protein in the context of continuum electrostatic calculations has been reiterated.[17–20] In many instances, single conformer pK_a calculations, typically performed on an X-ray structure, have failed to reproduce the experimentally observed ionization behavior of some residues. Both NMR-derived ensembles and MD-derived ensembles are thought to represent more realistically the molecular fluctuations in solution and their use for pK_a calculations has been suggested to overcome potential problems due to the fixed orientation of charges in a crystal structure.[9,12,21,22] The main limitation of MD and NMR ensembles, besides the computational or experimental costs, may be that both are fixed pH methods. They generate an ensemble of conformations representing the general fluctuations accessible to the protein around the fixed pH of the experiment or the simulation but do not take into account the pH dependent conformational variability. Therefore, their applicability to charged groups involved in coupled titration equilibria, where it is likely that a structural reorganization induced by the titration per se dominates over fixed protonation state fluctuations, may not be adequate.[16] However, if the fluctuations of titratable groups and permanent dipoles introduce the main uncertainty in the pK_a estimate, averages taken from both ensembles should improve accuracy. Which of the two ensembles meets better the needs for an accurate pK_a prediction is not clear. For the AB zipper and barnase, calculations on the NMR ensemble yield predictions comparable with (or sometimes better than) the results from taking averages from each single long (0.7–1 ns) trajectory. On the other hand, pK_a values of some residues are poorly predicted from NMR averages or from 1-ns MD runs and vice versa. Longer MD simulation time is required to narrow the gap between calculated and experimental pK_a for a few groups. However, if conformations generated by multiple, even short, MD simulations are considered, the MD ensemble improves the accuracy of the calculations significantly. It is likely that the success of a given ensemble in yielding accurate pK_a values is mainly governed by the extent of the conformational space that has been sampled. Previously, it was reported that 15 NMR conformers of ribonuclease H1 sample more conformational space than a 1.7-ns MD simulation.[58] This might be true for the NMR and MD ensemble of barnase as well. In contrast, the MD ensemble generated from multiple simulations might provide sampling of a broader conformational space than the NMR ensemble. We emphasize that with respect to pK_a calculations, the term "conformational space" must be used with care because local and global effects on the ionization behavior of a given group are not easily discernible. On the other hand, the pK_a property space is necessarily part of the conformational space. Therefore, the extent of conformational sampling will be critical in evaluating the merits of an ensemble of structures for pK_a calculations. However, because conformational sampling with the currently available methods cannot be truly exhaustive, one might ask how wide the sampling should be (without destroying the agreement with the experiment). Perhaps there is no general answer to this question, and the solution of the problem may be case oriented. As shown in this work, a correspondence with the experimental value is achieved with differing amounts of sampling for different residues. For some residues, sampling is not required at all, whereas for others multiple trajectories are required. A good indication that sampling is reasonably wide could be the presence of a finite number of discrete side-chain conformers interconverting on a reasonable timescale, as shown in Figure 4 for salt-bridge partners. In any case, effective sampling of thermally accessible conformations by multiple MD simulations has a great potential.

CONCLUSIONS

1. Conformational sampling provided by NMR ensembles and MD ensembles is an efficient way to account for the protein conformational flexibility in pK_a calculations in the framework of the continuum method. When sampling is sufficiently wide, the accuracy of pK_a predictions is higher at low protein dielectric constant. Moreover, pK_a values calculated from ensembles are less sensitive to the value of the protein dielectric constant.
2. Although single MD trajectories result in an improved overall accuracy, they can introduce unidirectional error for some sites. The results presented here indicate that the combination of multiple trajectories may help to overcome this problem. The length of simulation is not (the only) critical parameter for the quality of the overall prediction. Extended simulation times (1 ns instead of only 250 ps) are required to improve the final estimate for only a few groups. Because the calculation of MD trajectories is currently affordable for any reasonably sized protein, the use of multiple MD simulations

in the context of pK_a calculations deserves consideration.

3. For the proteins investigated in this work, titration curve averaging generally works better than conformational averaging over MD trajectories and over NMR conformers. Therefore, it appears that the overall improvement of pK_a predictions results from explicit conformational sampling rather than from the generation of a more representative structure.
4. Clustering of conformations is an efficient way to reduce the computational effort of pK_a calculations on few cluster representatives rather than hundreds of snapshot conformers. This does not reduce the informational content of MD-generated ensembles for the conformational flexibility, which is an important factor governing the electrostatic properties of proteins.

ACKNOWLEDGMENTS

We thank Drs. Andrey Karshikoff and James Warwicker for illuminating many important aspects of pK_a calculations and for stimulating discussions.

REFERENCES

1. Roxby R, Tanford C. Hydrogen ion titration curve of lysozyme in 6 M guanidine hydrochloride. Biochemistry 1971;10:3348–3352.
2. Tanford C, Roxby R. Interpretation of protein titration curves: application to lysozyme. Biochemistry 1972;11:2192–2198.
3. Kuramitsu S, Hamaguchi K. Analysis of the acid-base titration curve of hen lysozyme. J Biochem (Tokyo) 1980;87:1215–1219.
4. Bashford D, Karplus M. pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. Biochemistry 1990;29:10219–10225.
5. Yang AS, Gunner MR, Sampogna R, Sharp K, Honig B. On the calculation of pKas in proteins. Proteins 1993;15:252–265.
6. Antosiewicz J, McCammon JA, Gilson MK. Prediction of pH-dependent properties of proteins. J Mol Biol 1994;238:415–436.
7. Karshikoff A. A simple algorithm for the calculation of multiple site titration curves. Protein Eng 1995;8:243–248.
8. Warwicker J. Simplified methods for pKa and acid pH-dependent stability estimation in proteins: removing dielectric and counterion boundaries. Protein Sci 1999;8:418–425.
9. van Vlijmen HW, Schaefer M, Karplus M. Improving the accuracy of protein pKa calculations: conformational averaging versus the average structure. Proteins 1998;33:145–158.
10. Bashford D, Gerwert K. Electrostatic calculations of the pKa values of ionizable groups in bacteriorhodopsin. J Mol Biol 1992; 224:473–486.
11. Beroza P, Fredkin DR, Okamura MY, Feher G. Electrostatic calculations of amino acid titration and electron transfer, Q-AQB→QAQ-B, in the reaction center. Biophys J 1995;68:2233–2250.
12. Zhou HX, Vijayakumar M. Modeling of protein conformational fluctuations in pKa predictions. J Mol Biol 1997;267:1002–1011.
13. Demchuk E, Genick UK, Woo TT, Getzoff ED, Bashford D. Protonation states and pH titration in the photocycle of photoactive yellow protein. Biochemistry 2000;39:1100–1113.
14. Simonson T, Archontis G, Karplus M. A Poison-Boltzmann study of charge insertion in an enzyme active site: the effect of dielectric relaxation. J Phys Chem B 1999;103:6142–6156.
15. Schaefer M, Sommer M, Karplus M. pH-Dependence of protein stability: absolute electrostatic free energy differences between conformations. J Phys Chem B 1997;101:1663–1683.
16. Demchuk E, Wade RC. Improving the continuum dielectric approach to calculating pK_as of ionizable groups in proteins. J Phys Chem 1996;100:17373–17387.
17. Beroza P, Case DA. Calculations of proton-binding thermodynamics in proteins. Methods Enzymol 1998;295:170–189.
18. You TJ, Bashford D. Conformation and hydrogen ion titration of proteins: a continuum electrostatic model with conformational flexibility. Biophys J 1995;69:1721–1733.
19. Beroza P, Case D. Including side chain flexibility in continuum electrostatic calculation of protein titration. J Phys Chem 1996;100: 20156–20163.
20. Alexov EG, Gunner MR. Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. Biophys J 1997;72:2075–2093.
21. Antosiewicz J, McCammon JA, Gilson MK. The determinants of pKas in proteins. Biochemistry 1996;35:7819–7833.
22. Dillet V, Dyson HJ, Bashford D. Calculations of electrostatic interactions and pKas in the active site of Escherichia coli thioredoxin. Biochemistry 1998;37:10298–10306.
23. Ripoll DR, Vorobjev YN, Liwo A, Vila JA, Scheraga HA. Coupling between folding and ionization equilibria: effects of pH on the conformational preferences of polypeptides. J Mol Biol 1996;264: 770–783.
24. Wlodek ST, Antosiewicz J, McCammon JA. Prediction of titration properties of structures of a protein derived from molecular dynamics trajectories. Protein Sci 1997;6:373–382.
25. Marti DN, Jelesarov I, Bosshard HR. Interhelical ion pairing in coiled coils: solution structure of a heterodimeric leucine zipper and determination of pK(a) values of glu side chains. Biochemistry 2000;39:12804–12818.
26. O'Shea EK, Klemm JD, Kim PS, Alber T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 1991;254:539–544.
27. Baudet S, Janin J. Crystal structure of a barnase-d(GpC) complex at 1.9 A resolution. J Mol Biol 1991;219:123–132.
28. Buckle AM, Henrick K, Fersht AR. Crystal structural analysis of mutations in the hydrophobic cores of barnase. J Mol Biol 1993;234:847–860.
29. Bycroft M, Ludvigsen S, Fersht AR, Poulsen FM. Determination of the three-dimensional solution structure of barnase using nuclear magnetic resonance spectroscopy. Biochemistry 1991;30:8697–8701.
30. Brooks BR, Bruccorleri RE, Olafson BD, David JS, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. J Comput Chem 1983;4: 187–217.
31. MacKerell JAD, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, et al. Empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102:3586–3616.
32. Brunger AT, Karplus M. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. Proteins 1988;4:148–156.
33. Darden TA, York DM, Pedersen LJ. Particle mesh Ewald: an N.log(N) method for Ewald sums in large systems. J Chem Phys 1993;98:10098–11002.
34. Feller SE, Zang Y, Paster RW, Brooks RW. Constant pressure molecular dynamics simulation: the Langevin piston method. J Chem Phys 1995;103:4613–4621.
35. Rykaert JP, Ciccotti G, Berendesen 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.
36. Jorgensen WL, Charndrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
37. Caflisch A, Karplus M. Structural details of urea binding to barnase: a molecular dynamics analysis. Structure Fold Des 1999;7:477–488.
38. Daura X, van Gunsteren WF, Mark AE. Folding-unfolding thermodynamics of a beta-heptapeptide from equilibrium simulations. Proteins 1999;34:269–280.
39. Ferrara P, Apostolakis J, Caflisch A. Thermodynamics and kinetics of folding of two model peptides investigated by molecular dynamics simulations. J Phys Chem B 2000;104:5000–5010.
40. Bashford D. An object oriented programming suite for electrostatic effects in biological macromolecules. In: Ishikawa Y, Oldehoeft RR, Reynders, JVW, Tholburn M, editors. Scientific computing in object oriented parallel environments. Vol 1343 of lecture notes in computer science. Berlin: Springer; 1997. p 233–240.
41. Warwicker J, Watson HC. Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. J Mol Biol 1982;157: 671–679.
42. Beroza P, Fredkin DR, Okamura MY, Feher G. Protonation of interacting residues in a protein by a Monte Carlo method:

application to lysozyme and the photosynthetic reaction center of Rhodobacter sphaeroides. Proc Natl Acad Sci USA 1991;88:5804–5808.
43. Gilson MK. Multiple-site titration and molecular modeling: two rapid methods for computing energies and forces for ionizable groups in proteins. Proteins 1993;15:266–282.
44. Sitkoff D, Sharp KA, Honig B. Correlating solvation free energies and surface tensions of hydrocarbon solutes. Biophys Chem 1994;51:397–403.
45. Nozaki Y, Tanford C. Methods Enzymol 1967;11:715–734.
46. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
47. Caves LS, Evanseck JD, Karplus M. Locally accessible conformations of proteins: multiple molecular dynamics simulations of crambin. Protein Sci 1998;7:649–666.
48. Serrano L, Fersht AR. Capping and alpha-helix stability. Nature 1989;342:296–299.
49. Lumb KJ, Kim PS. Measurement of interhelical electrostatic interactions in the GCN4 leucine zipper. Science 1995;268:436–439.
50. Hendsch ZS, Tidor B. Electrostatic interactions in the GCN4 leucine zipper: substantial contributions arise from intramolecular interactions enhanced on binding. Protein Sci 1999;8:1381–1392.
51. Fairman R, Chao HG, Lavoie TB, Villafranca JJ, Matsueda GR, Novotny J. Design of heterotetrameric coiled coils: evidence for increased stabilization by Glu(−)-Lys(+) ion pair interactions. Biochemistry 1996;35:2824–2829.
52. Spek EJ, Bui AH, Lu M, Kallenbach NR. Surface salt bridges stabilize the GCN4 leucine zipper. Protein Sci 1998;7:2431–2437.
53. Kumar S, Nussinov R. Fluctuations between stabilizing and destabilizing electrostatic contributions of ion pairs in conformers of the c-Myc-Max leucine zipper. Proteins 2000;41:485–497.
54. Kumar S, Nussinov R. Salt bridge stability in monomeric proteins. J Mol Biol 1999;293:1241–1255.
55. Lumb KJ, Kim PS. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. Biochemistry 1995;34:8642–8648.
56. Oliveberg M, Arcus VL, Fersht AR. pKA values of carboxyl groups in the native and denatured states of barnase: the pKA values of the denatured state are on average 0.4 units lower than those of model compounds. Biochemistry 1995;34:9424–9433.
57. Warwicker J. Improving pKa calculations with consideration of hydration entropy. Protein Eng 1997;10:809–814.
58. Philippopoulos M, Lim C. Exploring the dynamic information content of a protein NMR structure: comparison of a molecular dynamics simulation with the NMR and X-ray structures of Escherichia coli ribonuclease HI. Proteins 1999;36:87–110.
59. Loewenthal R, Sancho J, Reinikainen T, Fersht AR. Long-range surface charge-charge interactions in proteins: comparison of experimental results with calculations from a theoretical method. J Mol Biol 1993;232:574–583.
60. Mossakowska DE, Nyberg K, Fersht AR. Kinetic characterization of the recombinant ribonuclease from Bacillus amyloliquefaciens (barnase) and investigation of key residues in catalysis by site-directed mutagenesis. Biochemistry 1989;28:3843–3850
61. McCammon JA, Harvey SC. Dynamics of proteins and nucleic acids. Cambridge: Cambridge University Press; 1987. AQ1: Please check page range.