Proteins in Vacuo. Denaturation of Highly-Charged Lysozyme Studied by Molecular Dynamics Simulations

C. T. Reimann,*,† I. Velázquez,‡ and O. Tapia‡

Division of Ion Physics, Department of Materials Science, Uppsala University, Box 534, S-751 21 Uppsala, Sweden, and Department of Physical Chemistry, Uppsala University, Box 532, S-751 21 Uppsala, Sweden

Received: June 8, 1998; In Final Form: August 19, 1998

Molecular dynamics (MD) simulations have been employed to address the structural changes of lysozyme (LYZ) in vacuo driven by Coulomb forces associated with multiple protonation, as occurs in electrospray ionization. Some general features which emerged from this MD study are that sufficiently high charge states drove significant unfolding and reshaping of the tertiary structure, while domain and secondary structures survived to some extent. Generally, higher charge was required to effect denaturation in comparison with experimental observations. Other physical factors not presently included in the model, viz., high internal energy, may be required to reduce the barrier for driving the unfolding process.

Introduction

Physical studies of gas-phase ionized biomolecules are motivated by the growing usage of matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS)[1] and electrospray ionization mass spectrometry (ESI-MS),[2] which both produce gas-phase ions of extremely massive biomolecules and polymers.[3,4] Recently, questions about the conformation of gas-phase (bio)polymers have been raised. One practical concern of current interest is the viability of noncovalent but nonetheless highly specific biomolecular complexes in the gas phase.[5,6] Several physical techniques are now employed to study the conformation of gas-phase protein and peptide ions at low structural resolution, viz., gas-phase hydrogen−deuterium exchange,[7] charge-stripping reactions,[8−12] collisional cross-section measurements,[13−17] energetic surface imprinting,[18−22] and kinetic energy release.[23−26] Results of applying these new techniques substantiate the idea that biomolecular ions occupy definite structures and manifest particular conformational changes.

Recent experimental works have focused on lysozyme (LYZ).[12,17,22] Gas-phase LYZ displays a number of unique conformers which can be manipulated in various ways; yet despite its complicated behavior it possesses an important simplification in that its polypeptide chain is selectively "pinned" in four places by disulfide bridges, possibly serving as a constraint on the unfolding pathways. Nonetheless, very little is known about the actual conformers manifested by experimental data, because of their low resolution. Computer simulations are required to supplement these data with information at atomic resolution, and such structural information is produced here with the help of molecular dynamics (MD) simulations. A successful interplay between experiment and theory may yield unique information on structural properties of biomolecules in vacuo.

Knowledge of the structure and fluctuation patterns for molecular ions in vacuo may be essential to better understand how their behavior in condensed matter environments depends on interactions intrinsic to the molecules themselves.[27] MD simulations of proteins and nucleic acids in water, using atom-resolved force fields, have been successful in reproducing most of the structure and fluctuation patterns deduced, for example, in high-resolution X-ray crystallography.[28−35] The implicit hypothesis in this class of simulations is that the intramolecular force field is not changed by the solvent. In the present paper, this hypothesis is retained insofar as the GROMOS force field is employed to simulate biomolecular ions in vacuo.

The main emphasis in the present work is on conformational evolution in response to global charging of a protein in vacuo. These simulations require that all long-range interactions be included in the calculations. As far as the authors are aware, MD in vacuo has not been reported for fully charged protein ions, though such procedures have been carried out for peptide ions[25,36] as well as for globally neutral proteins.[37] MD simulations of LYZ have been carried out in vacuo in order to assess the response of the structure of LYZ to the Coulomb repulsive forces associated with various charge states. The results of these simulations and comparisons with experimentally obtained information form the subject of this paper.

Models, Methods, and Analyses

Simulations. All MD simulations were carried out using the GROMOS-87 37C4 force field,[38] which is the parameter set used for condensed phase simulations. Conditions of constant volume and temperature were used. The initial velocities were taken from a Maxwellian distribution characterized by $T = 293$ K. During the first 5 ps, the simulated system was strongly coupled to a Berendsen thermal bath[39] with a relaxation time constant of 0.01 ps. Subsequently, the coupling was weakened by changing the relaxation time constant to 0.1 ps. The SHAKE algorithm allowed the use of a large MD time step, 2 fs. In calculating forces, nonbonded interactions were evaluated at every time step for all neighbors within 0.8 nm. Long-range Coulombic interactions between charge groups were evaluated for *all* charge group pairs, every 10 time steps. Periodic boundary conditions were not applied.

Seeding Structures. The atomic coordinates from the X-ray

* Corresponding author. Fax: +46-18-555-736. Tel: +46-18-471-3872. E-mail: curt.reimann@material.uu.se.
† Department of Materials Science.
‡ Department of Physical Chemistry.

10.1021/jp982543p CCC: $15.00
Published on Web 10/27/1998

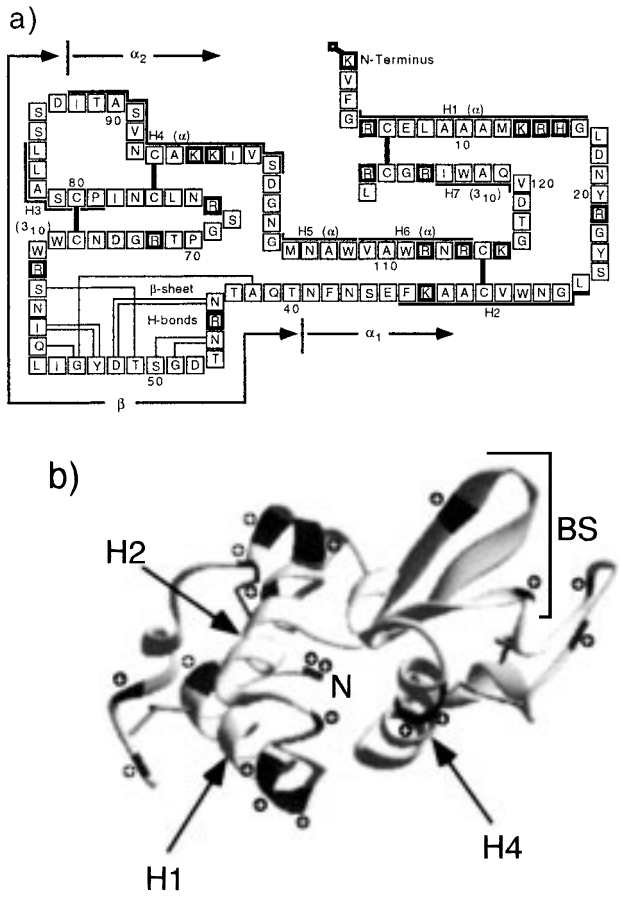

Figure 1. (a) Schematic "two-dimensional" diagram of the structure of hen-egg lysozyme. Heavy boxes denote residues whose side chains are charged at pH ≈ 3. Disulfide bridges are denoted by thick solid lines joining C residues. Helical elements and putative helix type are indicated for H1−7. Residues participating in the formation of the β-sheet (BS) are shown. Finally, residues participating in formation of (sub)domains α_1, α_2, and β are shown. (b) X-ray structure of LYZ visualized with Ribbons (M. Carson, University of Alabama at Birmingham, CMC). H1, H2, H4, BS, and the N-terminus are labeled. Charged residues are denoted by darkened sections of the polypeptide ribbon and are further emphasized by plus signs.

crystal structure of wild-type hen-egg-white LYZ at 0.17-nm resolution[40] were obtained from the Brookhaven protein data bank (E.C.3.2.1.17; PDB #1hel). Polar hydrogens were added and their coordinates generated from standard geometries. In total there were 1275 atoms in these simulations.

The X-ray structure was subjected to steepest-descent energy minimizations by molecular mechanics with bond-length constraints provided by SHAKE.[41] The total energy decreased by 7000 kJ·mol^{-1} (73 eV) after 350 iterations. This sterically relaxed structure was used to seed most of the MD simulations described below. To assess the response of an *already partially unfolded* LYZ conformer to applied charges, an unfolded conformer from a centrifugal denaturation MD trajectory[42] was used as a seeding structure in separate MD simulations.

Charging Model. To model fully charged LYZ, each basic amino acid residue (lysine, arginine, and histidine) side chain and the N-terminus were assigned a charge state of $q = 1+$ (Figure 1a), and all acidic amino acid side chains were kept neutral. This corresponds to the experimental situation in which LYZ is electrosprayed to the gas phase from a solution characterized by pH ≈ 3. These charge state assignments are typical of MD simulations carried out in the presence of H_2O and counterions, but in the present work these extra solvent particles were excluded, carrying out MD in vacuo.

A given charge state can be achieved in many ways.[43,44] Moreover, under some conditions such as high internal energy, adducted protons can be mobile.[45] However, since the aim of the present work was at a *statistical* treatment of Coulomb effects, a given desired charge Q was simply distributed over *all* the basic amino acid residues and the amino terminus (i.e., each q was set equal to Q/19), creating topology files for Q = 5, 9, 13, 16, and 19+.

The charge placement model described above was not constructed for the purpose of setting the dielectric constant: this parameter was set separately. A dielectric constant of $\epsilon = 2\epsilon_o$ (where ϵ_o is the dielectric permittivity of free space) is found to be appropriate for describing the gas-phase basicity of cytochrome c[11] and additionally accounts for maximum charge states observed in ESI of several other proteins.[43] However, the modeling used for obtaining ϵ is dependent on assumed protein conformation, and $\epsilon = 2\epsilon_o$ is a lower limit.[11,43] Also, dielectric constants greater than $2\epsilon_o$ are often used for quantum chemical calculations when large portions of the protein are modeled as a continuum.[46] The value $\epsilon = 2\epsilon_o$ was generally employed in the present work by dividing *all* charges in the model by $2^{1/2}$. Additional MD runs were performed for Q = 0+ while setting $\epsilon = \epsilon_o$ or $\epsilon = 2\epsilon_o$. Also, for $Q = 9$ and 19+, values of $\epsilon = \epsilon_o$ and $\epsilon = 2\epsilon_o$ were used. The results obtained were rather insensitive to this parameter.

Analyses. In the present studies, the focus was on global conformational changes. The MD output trajectories were analyzed mainly for energies, RMS position deviations from the X-ray structure, radii of gyration (R_{gyr}), and principle moments of inertia (PMI). The overall length of the protein was also calculated from the maximum distance between all pairs of atoms. Two domains were also investigated, α (residues 1−39 and 88−129) and β (residues 40−87).[47−50] The α domain was further considered to be composed of two subdomains, α_1 (residues 1−39) and α_2 (residues 88−129), which have been separately synthesized and studied.[50] These domains are shown in a two-dimensional map of the LYZ structure, along with a Ribbons diagram of the X-ray structure, in Figure 1.

The MD output trajectories were also analyzed for traces of secondary structure. The PDB entry lists seven helices, here labeled H1 through H7 and shown in Figure 1. H1, H2, H4, H5, and H6 are listed as α-helices, whereas H3 and H7 are listed as 3_{10} helices (see viz. ref 51). For helices initially present as α-helices, the distances between appropriate donor−acceptor centers were checked. A hydrogen bond was regarded to exist if this distance was smaller than 0.365 nm, a value set 0.05 nm in excess of the typical or representative hydrogen bonding distance of 0.315 nm[52] in order to register fluctuating hydrogen bonds. Plots of the number of hydrogen bonds surviving during each MD trajectory provided a detailed look at the α-helices; a more general view of them was provided by calculating and monitoring distances between C_α atoms at the ends of each helix; i.e., the end-to-end length was estimated for each helix.

The initial hydrogen bonding pattern of the β-sheet was determined from the X-ray structure, guided by a standard scheme for the β-sheet motif,[53] and is shown in Figure 1. The integrity of the β-sheet was determined by plotting the number of these hydrogen bonds surviving during each MD trajectory.

For some simulated structures, orientationally averaged collision cross sections appropriate for a helium buffer gas were estimated for comparison with experimental values derived in ion drift mobility experiments. Schemes for carrying out such calculations have been presented in the literature:[54,55] here an implementation of the hard-spheres projection approximation

was employed. Each conformer was placed successively in 120 orientations uniformly spaced on a tessellated[56] half-sphere. For each orientation, a projection of the atomic coordinates onto the plane truncating the half-sphere was carried out, and the hard-spheres collision cross section was calculated for normally incident He atoms. To do this, each protein atom was assumed to scatter any helium atom which would impact within a radius equal to the hard-spheres collision distance. Then the cross section is the total area consisting of all points which are within the hard-spheres collision distance of at least one protein atom; this was evaluated numerically with grid spacings of 0.01 nm along orthogonal directions in the projection plane. In the calculations, hard-spheres collision distances given in the literature[17] were employed except that since the united atoms scheme was used here, larger distances were required for the $-$CH, $-CH_2$, and $-CH_3$ united "atoms". For these units, collision distances of 0.30, 0.33, and 0.355 nm, respectively, were employed. After the collision cross sections for all the orientations were obtained, the orientationally averaged collision cross section for the protein conformer was computed. For the X-ray structure of LYZ, an orientationally averaged collision cross section of 1270 Å^2 was obtained, only 8% greater than the value 1180 Å^2 obtained by an analogous treatment.[17] The projection method ignores multiple scattering and thus provides lower limits for collision cross sections; true values can be ≈20% larger.[55] However, the focus here is on *relative* changes of collision cross sections with respect to that of the X-ray structure.

Results

Tertiary Structure. *(i) Fold Retention.* For charge states $Q \leq 13+$, the RMS deviations of the C_α atoms did not exceed ≈0.24 nm over the 1-ns integration time of the MD trajectories. This value indicates that a noticeable but still slight deviation from the X-ray structure occurred. This deviation reflected a minor compactation of the conformation, as also mirrored in a slight decrease of the radius of gyration during the MD trajectories. For $Q = 0$ and 5+, the overall length of the conformers decreased during the MD trajectories (Figure 2a, Table 1). However, the final overall lengths achieved for conformers with $Q = 9+$ and 13+ were slightly greater than that of the X-ray structure (Figure 2a, Table 1). Similar lengths computed as the maximum distance between any two C_α atoms in the conformer showed no such enhancement for $Q = 9+$ and 13+ (data not shown). Thus the polypeptide backbone was not significantly distorted up to $Q = 13+$, but some of the charge groups were repelled from the body of the conformer for $Q = 9+$ (Arg21 and Arg68) and $Q = 13+$ (Lys1, Arg14, Arg21, Arg45, Arg68, and Arg128).

During the trajectories for $Q \leq 13+$, somewhat more atoms came into energetically favorable steric interactions, since the Lennard-Jones energy decreased, as shown in Figure 2b. The global structure and time evolution of the LYZ was *generally* independent of charge state (the exception being $Q = 5+$, which displayed slightly lower RMS values). This is despite a clear general increase in the Coulomb potential energy with increasing charge state, as shown in Figure 2c. The conformer occurring at the end of each MD simulation is shown in Figure 3. The general trends mentioned above are recognized as a preservation of the globular shape along with a slight shortening of LYZ along its long axis.

(ii) Fold Disruption. For charge states $Q \geq 16+$, the RMS position deviations exceeded the ones observed for $Q \leq 13+$ after only 50 ps or less of MD integration. The RMS values

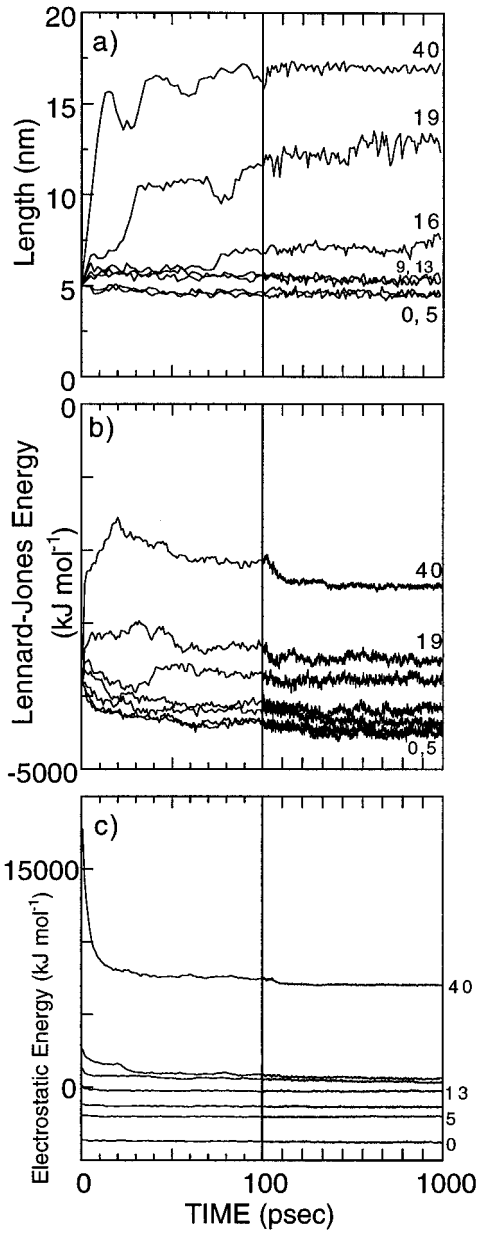

Figure 2. Plots of key indicators of overall structure of LYZ in various charge states with $\epsilon = 2\epsilon_o$. (a) Maximum length between any two atoms, roughly equivalent to the length along the protein long axis, (b) Lennard-Jones energy, and (c) electrostatic energy. To bring out details in the first 100 ps, the time scales are split into two segments. In (b) and (c), the final level of the curves increases in the order of increasing charge state.

displayed sizable oscillations, and, especially for $Q = 16$ and 19+, a multiple-plateau structure was observed, as also shown in the plot of conformer overall length (Figure 2a). The Lennard-Jones energy rose and then fell somewhat, reflecting that unusually high-energy steric contacts (motional frustration) occurred and were then relieved as the equilibrium close-packed interior of the protein was pulled apart over a structural energy barrier for unfolding (Figure 2b). As unfolding occurred, allowing charge groups to separate, greatly enhanced initial levels of electrostatic potential energy were reduced (Figure 2c). LYZ in charge state $Q = 19+$ achieved a length of around 13 nm (Table 1).

Domain and Subdomain Structure. *(i) Fold Retention.* For charge states $Q \leq 13+$, the (sub)domains α_1, α_2, and β all stayed near their crystallographic counterparts (Figure 4). In particular,

TABLE 1. Table of Lengths Derived for Some of the Gas-phase Protein Ion Species Studied by Molecular Dynamics[a]

species	length (starting MD from X-ray structure)	length (starting MD from partly unfolded structure)
neutral LYZ, $\epsilon = \epsilon_o$	4.4± (0.0) nm	
neutral LYZ	4.5 ± 0.1 nm	5.59 ± 0.1
LYZ^{5+}	4.5 ± 0.1 nm	
LYZ^{9+}, $\epsilon = \epsilon_o$	5.5 ± 0.1 nm	
LYZ^{9+}	5.5 ± 0.2 nm	7.26 ± 0.2
LYZ^{13+}	5.3 ± 0.2 nm	10.0 ± 0.3
LYZ^{16+}	7.5 ± 0.2 nm	11.2 ± 0.2
LYZ^{19+}, $\epsilon = \epsilon_o$	13.8 ± 0.2 nm	
LYZ^{19+}	13.0 ± 0.2 nm	

[a] Each length is the average length over the last 100 ps of a dynamics run. The standard deviations are given. The length is estimated as the maximum distance observed to separate two atoms of any type. Unless otherwise noted, $\epsilon = 2\epsilon_o$ was used. For comparison, crystal LYZ is characterized by a length of 5.0 nm.

the RMS deviations for α_1 hovered around 0.1−0.15 nm, while those for β were mostly between 0.1 and 0.2 nm (occasionally exceeding 0.2 nm on a transient basis). For α_2, the RMS deviations were greatest, hovering around 0.25 nm, indicating minor but perceptible conformational changes from the X-ray structure.

(ii) Fold Disruption. For charge state $Q = 16+$, the RMS deviations for α_1 showed a rise to a plateau of about 0.3 nm (Figure 4). After persisting for about 750 ps, the RMS deviation increased to above 0.4 nm, showing that an altogether different conformation was finally attained. The subdomain α_2 displayed an RMS deviation of around 0.3 nm for most of the simulation. For this charge state, β maintained a strong likeness to the X-ray structure, as its RMS deviations fluctuated around 0.2 nm for the whole simulated trajectory.

For $Q = 19+$, major alterations in the conformations of α_1 and α_2 occurred as both subdomains achieved RMS deviations as large as 1.1 nm. α_1 denatured most rapidly. In contrast to the α subdomains, β remained in a conformation similar to that of the X-ray structure, as its RMS deviations did not rise above 0.27 nm.

Some of the trends outlined above for the (sub)domains may be qualitatively observed in the structures shown in Figure 3. For example, with $Q = 16+$, it can be seen that the α_1 subdomain bulged away from the main body of the conformer, and the α_2 subdomain became slightly skewed as H5−6 separated from H4. By contrast, the β-domain remained quite compact and bore a passing visual resemblance to the corresponding portion of the X-ray structure.

Secondary Structure. Information about the fate of each helix and the β-sheet is given in Figure 5. Some of these results can be correlated against the visual appearance of the secondary structural elements shown in Figure 3.

For $Q \leq 13+$, secondary structures corresponding to the helices and β-sheet in the X-ray structure were well preserved. Indeed, Figure 5a shows that the end-to-end lengths of the helices were never far from 100%. Helices H1−4 and H6 were most invariant, while H5 and H7 manifested a significant shrinkage. In particular, the longest helices, H1, H2, and H4, were well preserved, with H4 exhibiting a slight but consistent lengthening to 110%. Figure 5b shows a similar degree of hydrogen bonding for H1 and H2; the degree of hydrogen bonding for H4 was sharply reduced, in line with the slight lengthening indicated in Figure 5a. The β-sheet showed a degree of survival comparable to that of the X-ray structure. Indeed, a moderately high charge state was correlated with an improvement in the viability of the hydrogen bonding pattern of the β-sheet.

For $Q = 16+$, see Figure 3, helix H1 was very strongly perturbed. This is confirmed by Figure 5a,b in that the end-to-end length of that section of the polypeptide chain was increased by about 150% and that the hydrogen bonding pattern characteristic of the α-helix abruptly disappeared. However, all other helical elements remained similar to the ones observed for lower charge states. In contrast, Figure 5b shows that the β-sheet presented a noticeable decrease in the number of surviving hydrogen bonds (to 25%), reflecting a conformational change of this secondary structure, that can be the reason for a slight rise in RMS on the β-domain for $Q = 16+$ (although the overall RMS value of 0.2 nm indicated in general a continuing close similarity to the X-ray structure).

For $Q = 19+$, see Figure 3, most of the helices became rather stretched, though helical motifs and nearby hairpin turns could still be perceived. Numerical evidence of what is observed in the figure is that H1 and H2 increased in length (Figure 5a) and only H4 had more than 40% surviving hydrogen bonds (Figure 5b). As seen in Figure 3, H4 was stretched, but mainly on the side toward the C-terminus past the disulfide bridge at residue 94. It should be noted that the β-sheet still presented 30% surviving hydrogen bonds (Figure 5b).

Role of Prior Unfolding. Unfolding of LYZ in vacuo produced by a weak centrifugal force has been determined recently by us.[42] Since in a real sample there always exist populations of folded and partially unfolded species, an attempt was made to assess the response of an *already partially unfolded* LYZ conformer[42] to an applied charge. Again, charge was smoothly distributed over all the basic sites. A picture of the response of the partially unfolded structure to Coulomb repulsion is obtained from a plot of RMS deviations of the C_α atoms with respect to the X-ray structure (Figure 6a) and plots of overall conformer lengths (Figure 6b).

If $Q = 0$, substantial reglobularization to a quite compact structure occurred, as indicated by R_{gyr} and conformer length values, though it was *far* from the X-ray structure as indicated by the RMS deviations. The striking point is that the persistence time manifested by the in vacuo MD trajectories extends at least 1 ns (data not shown). If $Q = 9+$, a modest reglobularization occurred, but to a conformation which was significantly further from the native one and significantly longer than the one reached by relaxation in $Q = 0$. The final structures documented for $Q = 0$ and $9+$ may be considered as belonging to a family of rather compact conformers. By contrast, using $Q = 13+$, the starting structure unfolded further to a final length of 10.0 nm, and using $Q = 16+$, the final length was 11.2 nm. These lengths exceeded by far the lengths reached when starting MD directly from the X-ray structure (Table 1). Moreover, the unfolded, charged species displayed an abundance of helical motifs, indicating that significant traces of secondary structure continued to be present.

Discussion

The experimental observation of specific, noncovalently bound complexes[5] and more elaborate biomolecular assemblies[57] intact as ions in vacuo suggests that native protein conformations can be stable against a certain degree of Coulomb repulsion. Yet ample experimental evidence also exists showing that protein ions can unfold in vacuo.[17] Here, results of MD simulations showed that the native conformation of lysozyme, a globular enzyme, was extremely robust with respect to

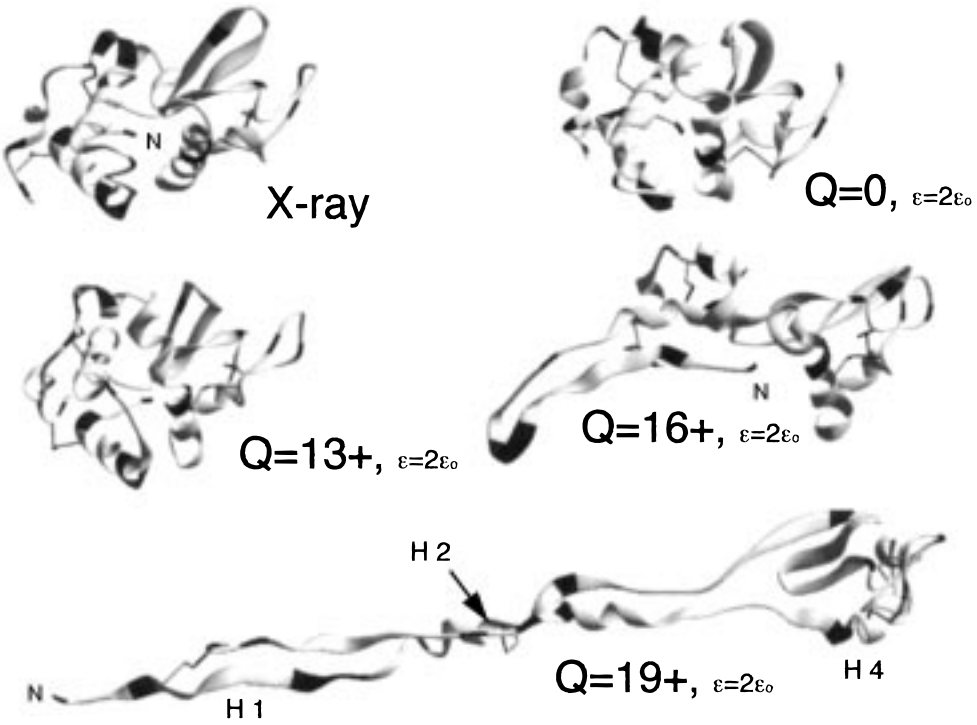

Figure 3. Structures of LYZ visualized after ≈1 ns of molecular dynamics simulation for the X-ray structure and for four different charge states. The N-terminus is labeled, and disulfide bridges are indicated. Charged residues are shown by darkened sections. Any helical or sheetlike structure was always presented in ribbon form, even after denaturation, to aid in identification of formerly intact secondary structural elements.

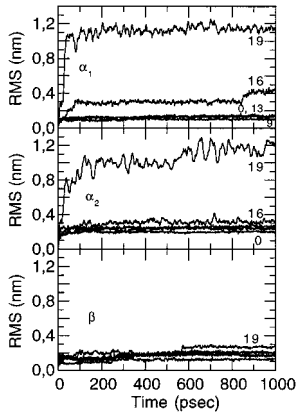

Figure 4. RMS deviations of the C_α atoms of the indicated (sub)-domains α_1, α_2, and β with respect to the same atoms in the X-ray structure. For the α_2 subdomain, the final level of the curves increases in the order of the charge state used. For the β-domain, only the last segment of the trace for $Q = 19+$, from 570 to 1000 ps, is specifically identified.

Coulomb-driven unfolding. MD also provided insight into the types of conformations adopted by lysozyme after the threshold for Coulomb-driven unfolding was reached.

Stability and Unfolding of Lysozyme. To initiate Coulomb-repulsion-driven unfolding, starting from the native LYZ structure and using the GROMOS force field, high charge states in excess of $Q = 13+$ were required. The X-ray structure of LYZ thus appears to be highly robust in vacuo, paralleling, viz., the known high degree of thermal stability of lysozyme in solution.[58]

Was the charge threshold for Coulomb-driven unfolding artificially increased because of the arbitrary, smooth application of charge to all the basic sites of the protein (scheme 1)? To investigate this question for $Q = 9+$, the charging scheme given by Miteva *et al.*[44] was employed (scheme 2) except that Lys96 (44% occupancy) was substituted for Asn77 (47% occupancy). An alternative charging scheme for $Q = 9+$ was to charge only the Arg side chains in the order given by Miteva *et al.*[44] (scheme 3).

For scheme 1, the $Q = 9+$ charge state manifested RMS position deviations $\leq$0.24 nm. However, for scheme 2, the RMS position deviations were $\leq$0.19 nm, and for scheme 3, these deviations were $\leq$0.21 nm. Therefore, for $Q = 9+$, charging specific side chains instead of smoothly distributing the charge over all the basic side chains certainly did *not* compromise the overall stability or the compact structure of the protein. This is also reflected in the fact that for all three charging schemes, a slight decrease of R_{gyr} and overall conformer length occurred during the MD trajectories, indicating a minor compactation of the conformation with respect to the X-ray structure. When comparing the three charging schemes, the potential energies at the end of the trajectories show that scheme 2 yielded the lowest energy conformation (-5450.2 kJ·mol^{-1}), whereas scheme 1 resulted in the highest energy conformation (-2414.1 kJ·mol^{-1}). Scheme 3 displayed an intermediate potential energy value for the final conformer (-3011.8 kJ·mol^{-1}).

For $Q = 9+$, the different charging schemes resulted in the same overall structure. Thus, the smooth distribution charging scheme provides useful general information in cases for which the specific charging scheme is not known. The different potential energy values listed above illustrate the fact that the smooth distribution charging scheme does induce a greater Coulomb repulsion. However, this charging scheme also weakens the ability of the charge groups to engage in self-solvation. Thus, charge groups which were oriented outward from the body of the conformer for $Q = 9+$ applied smoothly to all the basic residues as in scheme 1 were located closer to the body of the conformer when the total charge was applied as unit charges to a subset of the basic residues as in scheme 2.

Could the assumption of $\epsilon = 2\epsilon_o$ have drastically affected the outcome of the simulations? In the GROMOS force field,

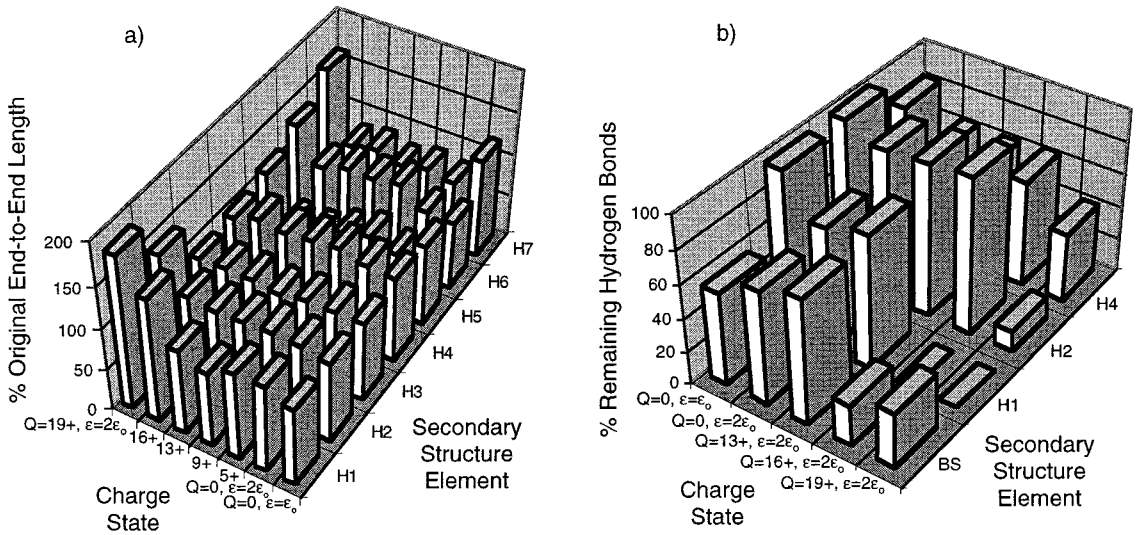

Figure 5. (a) Distances between C_α atoms at the beginning and end of each helical element in LYZ, averaged over the last 100 ps of each MD trajectory. These distances are expressed as a percentage of the same distance associated with the appropriate structural element in the X-ray structure. (For reference, completely stretched, linearized polypeptide chain is elongated by ≈240%.) For $Q > 0$, $\epsilon = 2\epsilon_o$ was used. (b) Number of surviving hydrogen bonds characteristic of several secondary structural forms, expressed as a percentage of the theoretical number of such bonds for native structures. Only the longest α-helices and the β-sheet are considered.

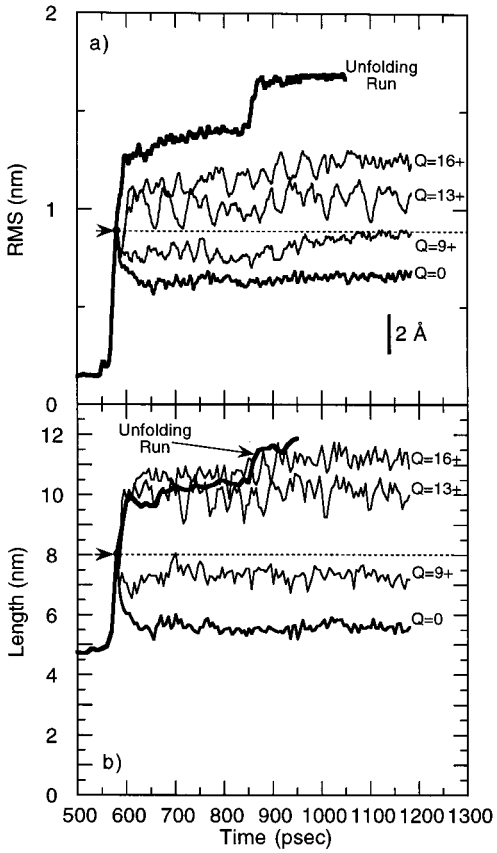

Figure 6. Global structure summaries of an unfolding trajectory[42] and relaxation MD trajectories for the unfolded conformer observed at 580 ps and charged to the charge states shown. In all runs summarized, $\epsilon = 2\epsilon_o$. The initial state is shown by a dot, an arrow, and a dotted line. (a) RMS deviations of the C_α atoms with respect to the X-ray structure; (b) overall lengths.

hydrogen bonding is represented by a combination of electrostatic and van der Waals interactions. In principle, by taking $\epsilon = 2\epsilon_o$ instead of $\epsilon = \epsilon_o$, weaker hydrogen bonding will result with possibly more open, "floppy" structures. In practice, additional MD runs carried out with $Q = 0$ showed that changing the dielectric constant did not result in any major deviation from the X-ray structure over the time of the simulation (1 ns). For $Q = 9+$, use of $\epsilon = \epsilon_o$ instead of $\epsilon = 2\epsilon_o$ did not produce unfolding; rather, a good resemblance to the X-ray structure was maintained. For $Q = 19+$, use of $\epsilon = \epsilon_o$ instead of $\epsilon = 2\epsilon_o$ resulted in a final conformer length only 3% greater.

The occurrence in the MD trajectories of an unfolding charge-state threshold between $Q = 13+$ and $16+$ is qualitatively consistent with estimates of free energy changes,[59] which suggest that compact, globular conformers become less stable than helices at charge states around $Q = 9$ or $10+$.[59] Estimates of free energy changes[59] also indicate that β-sheets are less stable with respect to Coulomb denaturation than helices, whereas in the present MD simulations partial preservation of the β-domain was observed, even for high charge states. This discrepancy most likely results from the fact that the simple model[59] does not account for the presence of intra-polypeptide-chain disulfide bridges, which provide additional stabilization at certain locations.

After integrating the trajectories for about 1 ns, the final conformer length reached was correlated with Q for $Q \geq 13+$ (Table 1). To determine an upper limit for the maximum extensibility of the LYZ protein under Coulomb repulsion, an artificially high charge state was used ($Q = 40+$), speeding up the sequence of events associated with Coulomb-driven unfolding. The $Q = 40+$ species underwent a rapid denaturation, achieving a length of 15.9 nm after ≈90 ps and maintaining this length until the end of the trajectory at 1 ns.

The structural results obtained from MD simulations show that, even when the tertiary structure was disturbed or even lost to a large extent, significant traces of the original (sub)domain and secondary structure persisted. One could venture that this relative stability of the secondary structure, when most of the protein is unfolded under nonphysiological conditions, is determined by the amino acid sequence itself. However, recent observations, viz., on the influence of context/environment on the formation of either an α-helix *or* a β-sheet from the *same* amino acid subsequence positioned in slightly different proteins[60] motivate a general reexamination of the factors influential

in determining the folding and relative stabilities of such secondary structures. For disulfide-intact LYZ, the radial nature of the Coulomb repulsion, combined with the stabilization afforded by the disulfide bridges, appeared to yield a specific pattern of sequential unfolding and partial loss of the secondary structure, where the secondary structure at the N- and C-termini of LYZ was preferentially destroyed, while some vestiges of the secondary structure of H2, H3, H4, H5, and the β-sheet remained.

Correlation with Experimental Results on Protein Ions in Vacuo. The native conformation was here adopted as the appropriate starting point to sense structural behavior in MD simulations of disulfide-intact LYZ in vacuo in various charge states produced by ESI. The X-ray structure was also considered to be an appropriate starting point for other theoretical studies addressing proteins in vacuo.[12,44] For LYZ, this assumption appears to be justified: prior to ESI, LYZ is typically stored in an acidic aqueous solution (pH $\approx$ 3) containing a solvent such as methanol, and in a similar environment, LYZ has been shown to be able to maintain its native conformation.[61] Therefore, at the instant LYZ comes into the gas phase, it occupies structures close to its native conformation, independent of charge states populated at pH $\approx$ 3.

The MD simulation results presented here can be compared to results or interpretations of three unique types of experiments/measurements: ion-mobility, gas-phase basicity, and energetic surface imprinting. These experiments provide conformational information at low resolution, necessitating that comparisons be of a qualitative nature.

(i) Compact Conformers. The results of gas-phase basicity experiments carried out on disulfide-intact LYZ[12] ions produced directly in the ESI process were interpreted as displaying sizable populations of conformations which are as compact, within errors, as the X-ray structure of LYZ. Moreover, in ion mobility measurements carried out on disulfide-intact LYZ,[17] charge states $Q = 5+$ to $10+$ display, under conditions of gentle collisional interaction with a bath gas, "highly folded" collision cross sections only $\approx$20% greater than the one calculated for the X-ray structure. The MD result presented above that LYZ in charge state $Q = 9+$ maintained a compact structure with close correspondence to the X-ray structure is consistent with these experimental results.

(ii) Partly Unfolded Conformers. LYZ in charge state $Q = 9+$ does not always maintain a compact structure. For example, in energetic surface imprinting experiments a length of 13.5 nm was inferred for this species.[22] Also, a partly unfolded species with $\approx$67% greater collision cross section than the X-ray structure can be generated by energetic collisions during injection of $Q = 9+$ disulfide-intact LYZ into the bath gas in ion mobility measurements,[17] and results from gas-phase basicity measurements[12] are also interpreted in terms of a partially unfolded conformer. It has been argued[22] that the elongated species observed by surface-imprinting may correspond to partly unfolded species inferred from gas-phase basicity and ion mobility measurements. In the present MD simulations starting from the X-ray structure, unfolding to a length of $\approx$13 nm could be achieved with charge state $Q = 19+$, and the unfolded species had collision cross sections enhanced by 60$-$80% ($\epsilon = 2\epsilon_o$ and ϵ_o, respectively) with respect to the X-ray structure. These figures are in qualitative agreement with information derived from experiments,[12,17] except that the charge required to drive these changes within the simulated time scale was too high. However, MD simulations starting from an already partly unfolded structure with $Q = 13+$ achieved a length of $\approx$10 nm and a collision cross section enhanced by $\approx$50%, quite compatible with experiments (note, $Q = 14+$ is achievable experimentally for disulfide-intact LYZ[62]).

(iii) Partial Retention of Secondary Structure. Conformer lengths derived from energetic surface imprint experiments,[21,22] and ion mobility measurements of collision cross sections,[16,17] indicate for many protein ions in vacuo considerably less than the fullest possible elongation of the polypeptide chain after unfolding. This admits retention of significant domain and/or secondary structure as a *general* hypothesis for unfolded protein ions in vacuo. This hypothesis is reinforced by the MD results presented above which displayed significant traces of surviving domain and secondary structure even after a charge threshold for Coulomb-driven unfolding was achieved (Figures 3 and 5). Moreover, an α-helix was recently shown experimentally and theoretically to be a stable secondary structural element of a peptide ion in vacuo.[25]

(iv) Correlation of Degree of Unfolding with Charge State. Collision cross section measurements[17] for highly folded LYZ in the charge state range $Q = 5$ to $10+$ display a subtle but discernible trend of increasing cross section with increasing charge state ($\approx$9% increase over this range). A similar experimental trend is observed for the bovine pancreatic trypsin inhibitor protein, which is also stabilized by disulfide bridges.[59] The MD results presented above display a very slight lengthening and opening up in the same charge state range ($\approx$9% increase in collision cross section), consistent with the experimental results. Proteins with no stabilizing disulfide bridges can display a more pronounced correlation between degree of unfolding and charge state, as demonstrated by energetic surface imprinting experiments with apomyoglobin and disulfide-*reduced* LYZ[21,22] and by collision cross section measurements for apomyoglobin and cytochrome *c*.[13,59] The MD results presented above treated an inherently more stabilized protein. Nonetheless, as soon as the barrier for unfolding was overcome, the MD results also displayed an enhanced correlation of degree of unfolding with charge state, in general compatibility with experiments.

(v) In MD, Too High Charge Is Required to Drive Unfolding from the Native Conformation. One possible explanation for this discrepancy is that experimental studies address the structures after time scales ≥ 1 μs, while MD simulations are typically limited to time scales $\leq$ 10 ns. After $\approx$1 ns, the time profiles of conformer length for $Q = 16$ and $19+$ displayed sizable fluctuations (Figure 2a), which suggest that further unfolding might occur eventually. Therefore, to obtain in simulations a degree of unfolding similar to what is experimentally observed, exaggerated charge states may be required in the model. This is analogous to the use of very high temperatures to drive unfolding in MD simulations.[63] Additionally, processes which may cause significant enhancements in internal energy, such as collisions with buffer gas, may reduce the charge threshold for Coulomb-driven unfolding. Such processes were not modeled in the MD simulations described above. However, we observed that $Q = 13+$ was sufficient to drive further unfolding of an *already partly unfolded* LYZ conformer (Figure 6). Since the Coulomb repulsion due to $Q = 13+$ was unable to initiate unfolding starting from the X-ray structure, it is concluded that the threshold charge state for Coulomb-driven denaturation was reduced if previous perturbations had driven partial unfolding.

(vi) Coulomb-Trapping of Partially Unfolded Structures. We also observed that a partially unfolded conformer with $Q = 9+$ reglobularized but to a conformation which was significantly

more open than the native one (Figure 6). This result shows that Coulomb repulsion can "trap" a "folding intermediate", as demonstrated experimentally.[17] In view of the universal effect of Coulomb repulsive forces and the independence of the force field from the identity of the protein, the results in Figure 6 for $Q = 0$ and 9+ taken together suggest generally that proteins which are electrosprayed to the gas phase in extended conformations may tend to collapse to more compact structures. Moreover, if a conformer held in an extended form by Coulomb repulsion is stripped of charge, significant reglobularization may subsequently occur. These more compact structures do not necessarily resemble the X-ray structure.

Conclusions

Molecular dynamics (MD) techniques have been employed to probe the stability and Coulomb-repulsion-driven denaturation of multiply charged disulfide-bond-intact lysozyme (LYZ) in vacuo. Lack of in vacuo experimental information at the atomic level precludes a detailed quantitative comparison between our simulation results and available experimental results. However, our simulation results are certainly analogous to the general observation that proteins in solution at lower and lower pH values — and thus higher charge states — finally become denatured. This correlation encourages a cautious comparison of our simulation results with available low-resolution experimental data in vacuo.

For low charge states, the simulated native structure of LYZ was extremely robust. However, high charge states drove significant unfolding of the tertiary structure, while preserving domain, subdomain, and secondary structures to some extent. Two similar states of LYZ — compact[12,17] and partly unfolded[12,17,22] — are manifested experimentally. The MD results suggest that the unfolded conformers observed experimentally can be highly elongated structures which are constrained by the disulfide bonds. Simulated conformers with lengths $\approx 3\times$ greater than that of the X-ray structure have collisional cross sections enhanced by only 50 to 80%, in qualitative agreement with experimental results.[17,22]

The use of MD combined with perturbation−relaxation[64] techniques allows an exploration of conformational space very far from the native conformation.[65,66] The results presented here, along with available experimental results on protein ions in vacuo, show that Coulomb repulsion can be a powerful force for driving unfolding transitions. Theoretical studies of protein unfolding will provide further detailed information on extended conformations which cannot yet be studied by classical techniques such as X-ray diffraction and nuclear magnetic resonance because of lack of a sufficiently dense sample.

Acknowledgment. We thank the Swedish Research Council for Engineering Sciences (TFR), the Swedish National Board for Industrial and Technical Development (NUTEK), and the Swedish Natural Sciences Council (NFR). We thank Plamen Demirev for drawing our attention to an important reference.[44]

References and Notes

(1) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299−2301.
(2) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64−71.
(3) Chen, R.; Cheng, X.; Mitchell, D. W.; Hofstadler, S. A.; Wu, Q.; Rockwood, A. L.; Sherman, M. G.; Smith, R. D. *Anal. Chem.* **1995**, *67*, 1159−1163.
(4) Smith, R. D.; Cheng, X.; Bruce, J. E.; Hofstadler, S. A.; Anderson, G. A. *Nature* **1994**, *369*, 137−139.
(5) Przybylski, M.; Glocker, M. O. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 806−826.
(6) Loo, J. A. *Bioconjugate Chem.* **1995**, *6*, 644−665.
(7) Wood, T. D.; Chorush, R. A.; Wampler, F. M.; Little, D. P.; O'Connor, P. B.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2451−2454.
(8) McLuckey, S. A.; van Berkel, G. J.; Glish, G. L. *J. Am. Chem. Soc.* **1990**, *112*, 5668−5670.
(9) Winger, B. E.; Light-Wahl, K. J.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 624−630.
(10) Loo, R. R. O.; Winger, B. E.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 1064−1071.
(11) Schnier, P. D.; Gross, D. S.; Williams, E. R. *J. Am. Chem. Soc.* **1995**, *117*, 6747−6757.
(12) Gross, D. S.; Schnier, P. D.; Rodriguez-Cruz, S. E.; Fagerquist, C. K.; Williams, E. R. *Proc. Natl. Acad. Sci.* **1996**, *93*, 3143−3148.
(13) Covey, T.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 616−623.
(14) Cox, K. A.; Julian, R. K.; Cooks, R. G.; Kaiser, R. E. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 127−136.
(15) von Helden, G.; Wyttenbach, T.; Bowers, M. T. *Science* **1995**, *267*, 1483−1485.
(16) Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1995**, *117*, 10141−10142.
(17) Valentine, S. J.; Anderson, J. G.; Ellington, A. D.; Clemmer, D. E. *J. Phys. Chem.* **1997**, *101*, 3891−3900.
(18) Reimann, C. T.; Quist, A. P.; Kopniczky, J.; Sundqvist, B. U. R.; Erlandsson, R.; Tengvall, P. *Nucl. Instrum. Methods Phys. Res. B* **1994**, *88*, 29−34.
(19) Quist, A. P.; Ahlbom, J.; Reimann, C. T.; Sundqvist, B. U. R. *Nucl. Instrum. Methods Phys. Res. B* **1994**, *88*, 164−169.
(20) Reimann, C. T.; Sullivan, P. A.; Türpitz, A.; Altmann, S.; Quist, A. P.; Bergman, A.; Oscarsson, S. O.; Sundqvist, B. U. R.; Håkansson, P. *Surf. Sci. Lett.* **1995**, *341*, L1019-L1024.
(21) Sullivan, P. A.; Reimann, C. T.; Axelsson, J.; Altmann, S.; Quist, A. P.; Sundqvist, B. U. R. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 329−341.
(22) Reimann, C. T.; Sullivan, P. A.; Axelsson, J.; Quist, A. P.; Altmann, S.; Roepstorff, P.; Velázquez, I.; Tapia, O. *J. Am. Chem. Soc.* **1998**, *120*, 7608−7616.
(23) Kaltashov, I. A.; Fenselau, C. C. *J. Am. Chem. Soc.* **1995**, *117*, 9906−9910.
(24) Adams, J.; Strobel, F. H.; Reiter, A.; Sullards, M. C. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 30−41.
(25) Kaltashov, I. A.; Fenselau, C. *Proteins: Structure, Function and Genetics* **1997**, *27*, 165−170.
(26) Szilágyi, Z.; Drahos, L.; Vékey, K. *J. Mass Spectrom.* **1997**, *32*, 689−696.
(27) Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2426−2427.
(28) *Computer simulation of biomolecular systems*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; ESCOM: Leiden, 1993; Vol. 2.
(29) Weerasinghe, S.; Smith, P. E.; Mohan, V.; Cheng, Y.-K.; Pettitt, B. M. *J. Am. Chem. Soc.* **1995**, *117*, 2147−2158.
(30) Cheatham, I. T. E.; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193−4194.
(31) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(32) Daura, X.; Hünenberger, P. H.; Mark, A. E.; Querol, E.; Avilés, F. X.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **1996**, *118*, 6285−6294.
(33) Daura, X.; Oliva, B.; Querol, E.; Avilés, F. X.; Tapia, O. *Proteins* **1996**, *25*, 89−103.
(34) Cheatham, T. E.; Kollman, P. A. *J. Mol. Biol.* **1996**, *259*, 434−444.
(35) Tapia, O.; Velázquez, I. *J. Am. Chem. Soc.* **1997**, *119*, 5934−5938.
(36) Wyttenbach, T.; van Helden, G.; Bowers, M. T. *J. Am. Chem. Soc.* **1996**, *118*, 8355−8364.
(37) Sanejouand, Y.-H.; Tapia, O. *J. Phys. Chem.* **1995**, *99*, 5698−5704.
(38) van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulation (GROMOS) Library Manual* Biomos, Nijenborgh 16, 9747 AG Groningen, The Netherlands, 1987.
(39) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(40) Wilson, K. P.; Malcolm, B. A.; Matthews, B. W. *J. Biol. Chem.* **1992**, *267*, 10842.
(41) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(42) Reimann, C. T.; Velázquez, I.; Tapia, O. *J. Phys. Chem. B* **1998**, *102*, 2277−2283.
(43) Schnier, P. D.; Gross, D. S.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1086−1097.
(44) Miteva, M.; Demirev, P. A.; Karshikoff, A. D. *J. Phys. Chem. B* **1997**, *101*, 9645−9650.
(45) Dongré, A.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118*, 8365−8374.

(46) Tapia, O. Quantum Theories of Solvent-Effect Representation: an Overview of Methods and Results. In *Intermolecular Forces* Ratajczak, H., Orville-Thomas, W.-J., Eds.; John Wiley & Sons Ltd.: Chichester, 1982; Vol. 3.

(47) Hünenberger, P. H.; Mark, A. E.; van Gunsteren, W. F. *Proteins: Strut., Funct. Gen.* **1995**, *21*, 196−213.

(48) McCammon, J. A.; Gelin, B. R.; Karplus, M.; Wolynes, P. G. *Nature* **1976**, *262*, 325−326.

(49) Brooks, B.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 4995−4999.

(50) Yang, J. J.; van den Berg, B.; Pitkeathly, M.; Smith, L. J.; Bolin, K. A.; Keiderling, T. A.; Redfield, C.; Dobson, C. M.; Radford, S. E. *Folding Design* **1996**, *1*, 473−484.

(51) Post, C. B.; Brooks, B. R.; Karplus, M.; Dobson, C. M.; Artymiuk, P. J.; Cheetham, J. C.; Phillips, D. C. *J. Mol. Biol.* **1986**, *190*, 455−479.

(52) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: Berlin, 1984.

(53) Bränden, C.; Tooze, J. *Introduction to Protein Structure*; Garland Publishing: New York, 1991.

(54) von Helden, G.; Hsu, M.-T.; Gotts, N.; Bowers, M. T. *J. Phys. Chem.* **1993**, *97*, 8182−8192.

(55) Shvartsburg, A. A.; Jarrold, M. F. *Chem. Phys. Lett.* **1996**, *261*, 86−91.

(56) Silla, E.; Tunon, I.; Pascual-Ahuir, J. *J. Comput. Chem.* **1991**, *12*, 1077−1088.

(57) Suizdak, G.; Bothner, B.; Yeager, M.; Brugidou, C.; Fauquet, C. M.; Hoey, K.; Chang, C.-M. *Chem. Biol.* **1996**, *3*, 45−48.

(58) Matthews, B. W. *Adv. Protein Chem.* **1995**, *46*, 249−278.

(59) Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119*, 2240−2248.

(60) Minor, D. L.; Kim, P. S. *Nature* **1996**, *380*, 730−734.

(61) Buck, M.; Radford, S. E.; Dobson, C. M. *Biochemistry* **1993**, *32*, 669−678.

(62) Loo, J. A.; Loo, R. R. O.; Udseth, H. R.; Fulton, J. L.; Smith, R. D. *Rap. Commun. Mass Spectrom.* **1992**, *6*, 159−165.

(63) Mark, A. E.; van Gunsteren, W. F. *Biochemistry* **1992**, *31*, 7745−7748.

(64) Soares, C. M.; Björkstén, J.; Tapia, O. *Mol. Sim.* **1995**, *15*, 35−46.

(65) Laidig, K. E.; Daggett, V. *Folding Design* **1996**, *1*, 335−346.

(66) Smith, L. J.; Fiebig, K. M.; Schwalbe, H.; Dobson, C. M. *Folding Design* **1996**, *1*, R95−R106.