Transitions in Chain Entanglement and Compactness Associated with in Vacuo Unfolding of Lysozyme Ions

Gustavo A. Arteca,*,†,‡ C. T. Reimann,§ and O. Tapia‡

Département de Chimie et Biochimie, Laurentian University, Sudbury, Ontario, Canada P3E 2C6, Department of Physical Chemistry, Uppsala University, Box 532, S-751 21 Uppsala, Sweden, and Department of Analytical Chemistry, Chemical Center, Lund University, Box 124, S-221 00 Lund, Sweden.

Received: October 16, 2000; In Final Form: March 22, 2001

Chain entanglement and compactness are two a priori independent properties that convey large-scale shape features of polymer conformations. In this work, we use these properties to monitor the *initial step* for the in vacuo unfolding of charged lysozyme ions. Using molecular dynamics simulations and a statistical model for the protein charge distribution, we show the existence of a narrow range of total charge within which compact (quasinative) and extended (partly unfolded) conformers can coexist. Denatured conformers are always found above the critical charge for unfolding, whereas quasinative structures are associated with low-charge states. We find that the global molecular shape of the accessible conformers is conserved over a range of temperatures, despite a shift in the critical charge as temperature increases. Within our model, in vacuo unfolding occurs for lysozyme ions with charges $+7$ and $+8$, when a thermal bath of $T = 500$ K is considered. This result is compatible with gas-phase experiments on lysozyme and suggests that unfolding results from the combined effect of heating and the Coulomb repulsion between charged residues.

Introduction

Advances in controlled desorption and electrospray ionization have recently made it possible to place "intact" biomolecular ions in a controlled manner into a gas phase or a vacuum.[1−5] As a result, some conformational properties of unsolvated biomolecules (e.g., anhydrous proteins) can now be probed in flight, using experimental techniques such as mass spectrometry,[6] gas-phase basicities and proton-transfer rates[1], H/D exchange rates,[4] ion drift mobility,[1,5] and high-energy surface imprinting.[7,8] The structural data generated from these sources is typically of low resolution (e.g., mean molecular sizes or cross-sectional areas). Nevetheless, the experiments do provide sufficient information to derive a qualitative picture of the configurational transitions accessible to a fully dehydrated protein. Anhydrous protein ions are interesting in their own right, exhibiting characteristic behaviors that are not yet fully understood.[5] Moreover, these systems can teach us a great deal on the role of solvent in protein folding,[9−11] given that all internal rearrangements in unsolvated proteins are determined by properties intrinsic to the primary sequence.

Experiments on protonated proteins in the gas phase show inequivocally that both unfolding and refolding processes can take place even in absence of water.[5] The possibility of a folding transition in vacuo was already conjectured in the literature;[9] recent results put some of these notions on firmer ground, as well as rule out others. Given that the structural characterization of biomolecular ions is still in its infancy, it is not surprising that the impact of some recent developments has not yet been fully appreciated. To dispel any misconceptions, we summarize briefly here a number of key observations emerging from studies in the gas phase. The most revealing results have been derived from measurements of diffusion cross sections and gas-phase basicities of electrosprayed ions of proteins such as lysozyme, cytochrome *c*, BPTI, and apomyoglobin.[1−5] These experiments show the following: (a) Low-charge protein ions electrosprayed from a solution containing the folded (native) protein appear only in compact conformations. Their mean sizes, shapes, and reactivities with gas-phase bases are indistinguishable from the values observed for the native protein in solution. (b) Protein ions electrosprayed from a solution containing the acid-denatured protein appear only in noncompact (i.e., unfolded) conformations. (c) When the unfolded ions are exposed to a base in the gas phase, proton stripping can produce species with lower total charge. If the charge state is sufficiently low, the new ions return to compact conformations, with cross sections and reactivities indistinguishable from those observed in ions electrosprayed under folding conditions. (d) Cross sections change depending on the injection voltage and temperature of the gas cell chamber. Collisional heating with the buffer gas in the drift tube appears to facilitate unfolding. (e) Data indicate clearly that all configurational rearrangements take place in fully desolvated ions. The factors driving these transformations appear to be intrinsic to the dry protein. (f) The occurrence of compact ions with significant total charge cannot be reconciled with a conjectured "inside-out" fold in vacuo.[9] Given that protonation takes place on polar residues, Coulomb repulsion prevents burying them into a compact core.[2]

Altogether, the above evidence is consistent with the notion that, under electrospray ionization conditions, gas-phase ions can conserve some "memory" of their initial configurations in solution. Moreover, the results indicate the occurrence of solvent-free refolding. So far, it is not known whether gas-phase refolding produces compact conformers that resemble the native structure in solution. However, data does not rule out the possibility that the native state be at least *metastable* in the gas

* To whom correspondence should be addressed. Permanent address: Laurentian University. E-mail: gustavo@nickel.laurentian.ca

† Laurentian University.

‡ Uppsala University.

§ Lund University.

10.1021/jp0037955 CCC: $20.00
Published on Web 05/09/2001

phase, as well as the possibility of refolding into quasinative shapes without the intervention of water. By combining computer simulations of protein ion dynamics together with this body of observations, we are now in position to gain first insights on how biomolecular folding processes take place in vacuo. In this work, we consider a particularly well-defined aspect of the problem, namely, the nature of the *very early unfolding steps* for anhydrous protein ions in variable charge. Using nanosecond-long molecular dynamics (MD) simulations of lysozyme ions, we discuss some general properties of the accessible configurational space and the conditions that control the onset of in vacuo unfolding.

Our goal is to set up a simulation protocol that is consistent with available quantitative data. On the one hand, it appears that protein ions can conserve the native fold over a range of charged states.[5,12,13] Another important piece of information is the occurrence of a narrow range of values in the total protein charge, $q > 0$, wherein folded *and* partly unfolded conformers can coexist.[3−5] We refer to this as the *critical charge* for unfolding, q^*. Fully denatured conformations appear as *high-charge* states, where $q > q^* + 1$. In contrast, *low-charge* states, where $q < q^* - 1$, remain in compact quasinative folds. Under the experimental conditions of electrospray from acidic solutions, a typical value for the critical charge is $q^* \approx 9 \pm 1$ in lysozyme.[2,3] Below, we investigate the origin of this q^* value, as well as discuss the nature of the conformers generated near the unfolding transition as a function of q.

Recently, the unfolding of lysozyme at room temperature ($T = 293$ K) has been studied with a simple "statistical" model for the distribution of hydrogen ions on the protein.[14,15] In this approach, *static* charged states are generated by interpolating between the neutral and fully charged lysozyme at pH ≈ 3 (where *max* $q = 19$). (This procedure is qualitatively similar to averaging over several static proton distributions in more realistic models.[14] For clarity, we refer to the total charge in the statistical approach as q_s, whereas q is reserved for the experimental value.) Computer simulations with this model[14,15] show that unfolding does take place over a narrow range of values in the total charge, $q_s^* \pm 1$. However, the critical charge simulated at $T = 293$ K ($q_s^* \approx 15$) is much larger than the experimental value for q^*.[2,3] There is reason to believe that this difference is due to the fact that, under experimental conditions, the ions are not found at room temperature. First, ion-drift mobility experiments show that the q^* value decreases when the ions enter the drift tube with a higher injection energy. A higher injection energy must correlate with increased collisional heating with the bath gas.[5] Also, collision cross sections and MD simulations on cytochrome *c* show that a $q = 7$ ion adopts quasinative at 300 K, yet is fully unfolded at 573 K, whereas a $q = 5$ remains folded at both temperatures.[16] Below, we test this notion by studying the temperature effect on q^* for the initial step in the unfolding of lysozyme ions.

From a more fundamental viewpoint, our goal is to produce a survey of the configurational subspace traversed during in vacuo unfolding. To this end, we focus on the pattern of large-scale molecular shape properties that can distinguish between the various conformational changes associated with folding−unfolding transitions. These rearrangements include polymer collapse, loop diffusion, and the organization of a secondary structure.[9−11] As a convenient approach, we use the backbone anisometry and entanglement complexity to generate a two-dimensional "molecular shape space" for protein conformers.[17] Whereas the anisometry conveys the extent of chain collapse,[18] the degree of chain entanglement measures differences in topology[19−21] and discriminates among tertiary folds[22] with the same mean size. Recently, we have used this method to monitor the conformers found during in vacuo relaxation dynamics of unfolded *neutral* lysozyme, with and without disulfide bridges.[23,24] Here, we apply this technique to study unfolding transitions in *charged* lysozyme. Using a large manifold of MD runs, we estimate the mean molecular shapes for protein ions as a function of total charge and temperature. From the practical viewpoint, our work serves to test the reliability of MD simulations to model the experimental behavior of in vacuo lysozyme. From a more general viewpoint, our results uncover a simple pattern of accessible conformations, common to the ensemble of MD trajectories initiated under different initial conditions of total charge, temperature, and velocities.

Molecular Dynamics Simulation of Unfolding in Lysozyme Ions. The main factor triggering the in vacuo unfolding of protein ions is believed to be the Coulomb repulsion between charged amino acids. This process is at work even at low temperature, and it is akin to the unfolding of a protein in acidic solutions. High temperature can also cause unfolding, either by facilitating the Coulombic denaturation of the charged protein or by opening entirely distinct unfolding pathways. MD simulations have recently been used to study the unfolding of in vacuo lysozyme, both neutral[26,27] and charged.[14,15] In all cases, simulations produce elongated conformers, whose lengths and cross sections are compatible with the experimental glimpses of the unfolded state.[3,8]

In this work, we follow the protocol used in ref 14 to model lysozyme ions. The initial structure for all MD simulations is the (energy-minimized) native form of hen egg-white lysozyme (PDB code 1hel, 129 amino acids), with *intact* disulfide bridges. Trajectories are generated with the GROMOS87 37C4 force field.[28] Despite having been designed for simulations in condensed phase, this potential-energy function exhibits several features that make it attractive to study anhydrous proteins. First, the intraprotein parameters in the force field make no explicit reference to the solvent, nor do they intend to represent "implicit solvent" information. (If one were interested in simulating a *solvated* protein, the interaction of explicit solvent molecules with the protein would be described by additional potential-energy parameters without affecting the intraprotein force field. In our case, we deal with a solvent-free protein.) Second, the GROMOS force field contains only *generic* potential-energy terms, including stretchings, bendings, torsions, and van der Waals terms deduced for atoms in vacuo. In the absence of an alternative potential energy function *explicitly* parametrized with gas-phase-only data, our present approach represents a pragmatical choice. Other choices may eventually become available as soon as potentials are developed that are rigorously transferable from condensed phase to vacuum.

In addition, we have employed the GROMOS87 37C4 force field *to handle in vacuo boundary conditions*. The latter imply that the system can freely rotate and translate after equilibration. Neither periodic boundary conditions nor cutoffs for Coulomb interactions are used. Given that partial charges are considered explicitly, the dielectric constant is taken as $D = \epsilon/\epsilon_0 = 1$. Initial velocities are chosen from a Maxwell−Boltzmann distribution at the desired working temperature T. During the trajectory, the temperature is maintained through a weak coupling ($\tau = 0.1$ ps) to a Berendsen thermostat.[29] For a given protein model, we have generated independent MD runs by changing the initial distribution of velocities consistent with a bath temperature T. It should be noted that we are interested in the configurational rearrangements that take place during the initial steps of

unfolding. In other words, we monitor nonequilibrium transitions. The Berendsen thermostat may not be appropriate for computing canonically averaged properties;[29] yet, it is an common (albeit "uncontrolled") tool for introducing a dissipation mechanism for unfolding transitions. Because of the limitations of a MD approach with a Berendsen thermostat, we base our conclusions not on single events but rather on the common pattern emerging from a number of distinct MD trajectories.

To perturb the native fold, the protein must be charged. Charge states were created with respect to a nominal maximum charge achieved at pH $\sim$ 3. In lysozyme, this corresponds to a $q = 19$ ion, where eleven arginines, six lysines, one histidine, and the N terminus are protonated. Lower charge ions were modeled by using what we loosely call a "statistical approach," whereby the individual basic sites share an equal *fraction* $q_s/19$ of the total charge.[14,30] The resulting model has nonphysical fractional protonation sites. However, it can be viewed as an *effective* model mimicking the average behavior emerging from MD simulations of protein ions in all possible static distributions of a total charge q (i.e., those with $+1$ charges shuffled among 19 residues, up to a maximum of $+q$). Note that the latter "realistic" protein ions, and protons, are assigned to the same specific basic residues throughout the trajectory. Consequently, this approach ignores the possibility of proton mobility, a physical process believed to occur during the experiment.[5] Our present statistical model provides a partial, very simplified solution to this shortcoming. It must be clearly noted that we do *not* attempt to simulate the dynamics of proton exchange, because variations in protein charge would take place in a time scale comparable to large-scale intramolecular motions. The so-called "statistical charge model" does not represent implicitly or explicitly proton dynamics, nor should it be conceived as route to incorporate a potential of mean force for proton motion into the MD simulations. Rather, our model should be viewed as an estimate of the mean behavior that would emerge from a number of independent simulations of *static charge*, each one starting from a different initial charge state (both in terms of its total charge and distribution of individual charged sites). Results in ref 14 for several static-charge "realistic" lysozyme ions suggest that this estimate is, at least qualitatively, reasonable. Otherwise, the statistical charge scheme is not commensurate to MD simulations including a mean field for proton transfers.

Within the statistical model, lysozyme unfolds smoothly for $q_s \geq 16$ at $T = 293$ K[14] and at lower charges for higher temperatures.[15] Below, we compare the unfolding behavior of protein ions at high and low temperatures. In particular, we monitor the evolution of large-scale molecular shape features along the unfolding transition, seeking for properties that may be common to protein ions in various conditions. To this end, we use a series of test MD runs, each one lasting up to 3 ns. As we show below, the results from these runs suggest that 3 ns can be a sufficiently long period to uncover the *early stages* of a process as fast as the in vacuo unfolding by Coulomb repulsion. It should be noted that, within the experimental setup for ion-drift mobility measurements, proteins ions arrive at the detector after ca. 1 ms after the injection into the drift tube. However, it is believed that most of this time is simply used to migrate within a cool bath gas, long after the configurational transitions have taken place following the injection step. Thus, unfolding and refolding in vacuo would appear to take place much faster than in solution. Because we are interested only in the initial pattern of unfolding, our simulations are consistent with a setup implying a short (nanosecond-long) heating step.

Large-Scale Molecular Shape Features. Because the unfolding of lysozyme produces elongated conformers, we expect an increase of anisometry (and a loss of polymer compactness along the transition). These changes can be measured in terms of the chain asphericity Ω:

$$\Omega = \frac{(I_2 - I_1)^2 + (I_3 - I_2)^2 + (I_3 - I_1)^2}{2(I_3 + I_2 + I_1)^2} \qquad (1)$$

expressed as a function of the principal moments of inertia $\{I_i\}$.[18] As the protein evolves from a compact spheroid to a loosely packed hairpin, we expect a change from $\Omega = 0$ to $^1/_4$. However, the tertiary fold may still change at constant asphericity (e.g., by loop diffusion). To capture this aspect of polymer shape, we monitor the evolution of the self-entanglement complexity (i.e., the interpenetration of chain loops).[17,22] This property can be measured in terms of the *mean overcrossing number* $\bar{N}$, defined as the number of bond−bond projected crossings ("overcrossings") averaged over all spatial directions. By sampling m projections, and collecting the m_N ones with N overcrossings, we have[22]

$$\bar{N} = \lim_{m\to\infty} \sum_{N=0}^{\max N} N\frac{m_N}{m} \qquad (2)$$

where max $N = (n - 2)(n - 3)/2$ *for an* n-monomer chain. This descriptor includes information on polymer geometry *and* connectivity.[31,32] In biopolymers with permanent entanglements (e.g., DNA knots), the parameter $\bar{N}$ correlates closely with topological complexity.[19−21] In our case, chain self-entanglements are only transient, and $\bar{N}$ measures their fluctuations and persistence over time. As a chain disentangles by reorganizing (or eliminating) secondary structural elements, $\bar{N}$ decreases, reaching the formal limit $\bar{N} \to 0$ in a molecular rod. The two-dimensional $(\bar{N},\Omega)$ map provides a simple space to monitor the evolution of important molecular shape features.[23] Here, we follow the changes in Ω and $\bar{N}$ along MD trajectories of lysozyme ions at high temperature.

Temperature Effect on the Structural Stability of Lysozyme Ions. Fluctuations of the native fold center around $\bar{N} = 45.2 \pm 0.1$ and $\Omega = 0.0356 \pm 0.0002$. The molecular shapes for *low-* and *high-charge* states can be compared to these values. In low-charge states, ions remain in quasinative folds during all MD runs (i.e., they fluctuate among structures whose molecular shapes have $\bar{N}$ and Ω values similar to those of the native state). In contrast, ions in high-charge states unfold rapidly and do not produce partly unfolded transients. At room temperature, the native fold is conserved in neutral lysozyme[25,26] and in lysozyme ions with $q_s < 14$.[14] Recent simulations indicate that temperature can affect the onset of unfolding in protein ions.[15] Here, we shall discuss in more detail the dependence of the critical charge q_s* with the temperature and characterize the configurational transitions that take place for $q_s > q_s$*.

Figure 1 illustrates the general effect of temperature on the unfolding of a $q_s = 9$ ion by displaying four different MD trajectories as paths in molecular shape space. All trajectories start from the native conformation and last 3 ns; except for the temperature of the simulated bath, all initial conditions are the same. This $(\bar{N},\Omega)$ map is representative of the behavior observed at other charges and temperatures. For $q_s = 9$, we observe:

(i) At room temperature, $q_s = 9$ represents a low-charge state. Here, the ion fluctuates among conformations which are only slightly distorted with respect to the native structure. These conformers are more compact than the fold in the crystal but

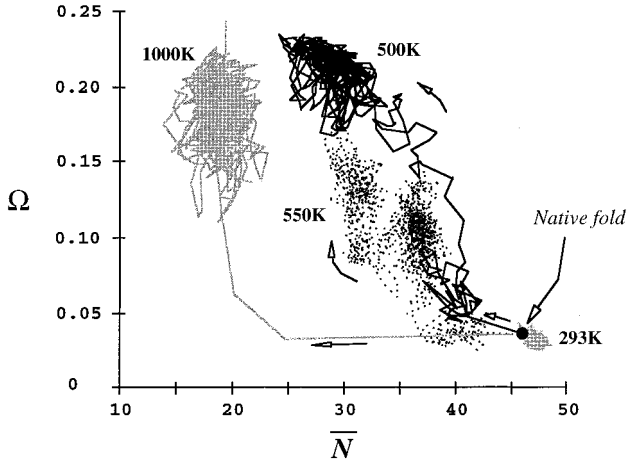

Figure 1. Molecular shape map for the in vacuo unfolding dynamics of lysozyme ion $q_s = 9$ at various temperatures. All simulations start from the same conformer with native fold. At room temperature, $q_s = 9$ represents a "low-charge" state where the ion does not unfold. At temperatures between 500 and 550 K, $q_s = 9$ represents a "high-charge" state leading to unfolding. At $T = 1000$ K, relaxation produces a fast "melt-down" of the secondary structure (i.e., a decrease in $\bar{N}$), followed by an elongation of the chain by Coulombic repulsion (i.e., an increase in Ω).

less compact than the in vacuo native structure for neutral lysozyme.[14] This result is due to a balance between two opposing effects: a "structural constriction" caused by a lack of solvent electrostatic screening in vacuo[5c] and a "moderate swelling" due to the repulsion among charged amino acids in the $q_s = 9$ state.

(ii) Between 500 and 550 K, the native structure is unstable; therefore, $q_s = 9$ represents a high-charge state. The associated unfolding paths show a gradual decrease in entanglement concurrent with an increase in anisometry. This behavior is consistent with a gradual *concerted configurational rearrangement* along the trajectory, one where loops slowly diffuse apart as the secondary structure unfolds.

(iii) At high temperature (e.g., $T = 1000$ K in Figure 1), the $q_s = 9$ ion unfolds in two steps: first, we observe a fast "melt-down" of secondary structure (represented as a decrease in $\bar{N}$), followed by a repulsion of charged loops (seen as an increase in Ω). The result is a rapid transformation of the native fold into a random chain with strong repulsive interactions. A similar behavior is observed in highly charged lysozyme at low temperature.

An effective temperature of 500 K is not inconsistent with the condition of in-flight protein ions under experimental setups with high injection energy. Thus, our results appear to agree with experiments that indicate that lysozyme ions are already denatured for $q \approx 9$.[3] In addition, the above behavior resembles qualitatively the temperature effects recently observed in the unfolding of cytochrome *c* ions.[16] In the next section, we study in more detail this regime of conformational rearrangements and discuss how the unfolding of lysozyme ions at $T = 500$ K compares with the transition at room temperature.

Molecular Shape Transitions and Critical Charge for Unfolding at 500 K. We have studied lysozyme ions at 500 K by performing a series of MD simulations in various charge states. Critical trajectories were repeated with different initial velocities to assess variability and estimate mean conformational shapes.

Examples of the dynamic behavior at 500 K are given in Figure 2. The figure shows the evolution of the mean over-

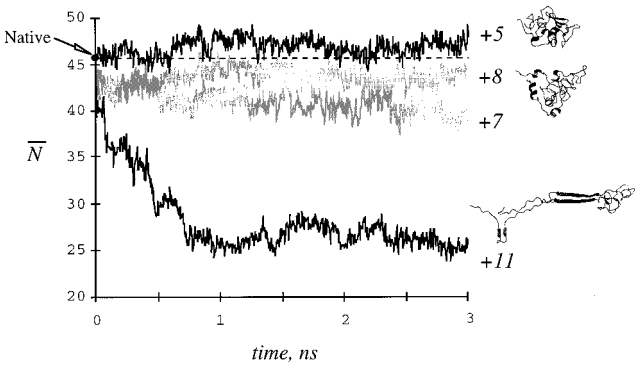

Figure 2. Changes in backbone entanglement for lysozyme at 500 K. The lines are single MD trajectories generated with identical initial conditions but different charged states, $q_s = 5$, 7, 8, and 11. The dashed segment stands for the $\bar{N}$ value characteristic of native (neutral) lysozyme in the crystal structure. Snapshots illustrate representative conformers observed at the end of the trajectories for $q_s = 5$, 8, and 11. There is full unfolding by Coulombic repulsion for $q_s = 11$.

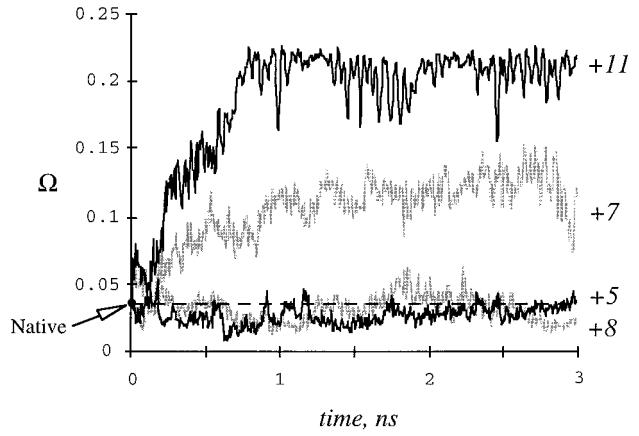

Figure 3. Changes in backbone anisometry for lysozyme at 500 K. The results for $q_s = 5$ and 8 are similar in asphericity, yet, are clearly distinguished in terms of entanglements (cf. Figure 1).

crossing number in single MD runs for lysozyme ions with charges $q_s = 5$, 7, 8, and 11. All simulations start from the native conformer and share exactly the same initial distribution of velocities for $T = 500$ K. The snapshots at the end of the trajectories correspond to representative conformers at 3 ns. (Images were generated with the program *MolMol*.[33]) Three qualitatively different results are observed: (a) Chain entanglements can be stable along a trajectory, reaching slightly higher $\bar{N}$ values than the native fold at $q_s = 5$ and slightly $\bar{N}$ lower values at $q_s = 8$. (b) Entanglements can remain stable, although significantly smaller than the native fold (e.g., $q_s = 7$). (c) A chain can disentangle with extensive fluctuations in $\bar{N}$ (e.g., $q_s = 11$). The conformational snapshots suggest that these three situations correspond, in terms of global folding features, to the following possible outcomes: (a) retaining a quasinative fold, (b) producing partly unfolded intermediates, and (c) reaching totally unfolded (elongated) conformers, respectively.

Figure 3 complements the information in Figure 2 by displaying the evolution of chain asphericity along the same MD runs. In the cases of $q_s = 7$ and 11, it is clear that the initially compact native structure has been strongly distorted. (The quasiregular oscillations in Ω for $q_s = 11$ correspond to large-scale bending of "hairpin" conformers as the one displayed in Figure 2.) In contrast, the lysozyme ions with charges $q_s = 5$ and 8 have similar anisometries, although their chains can be

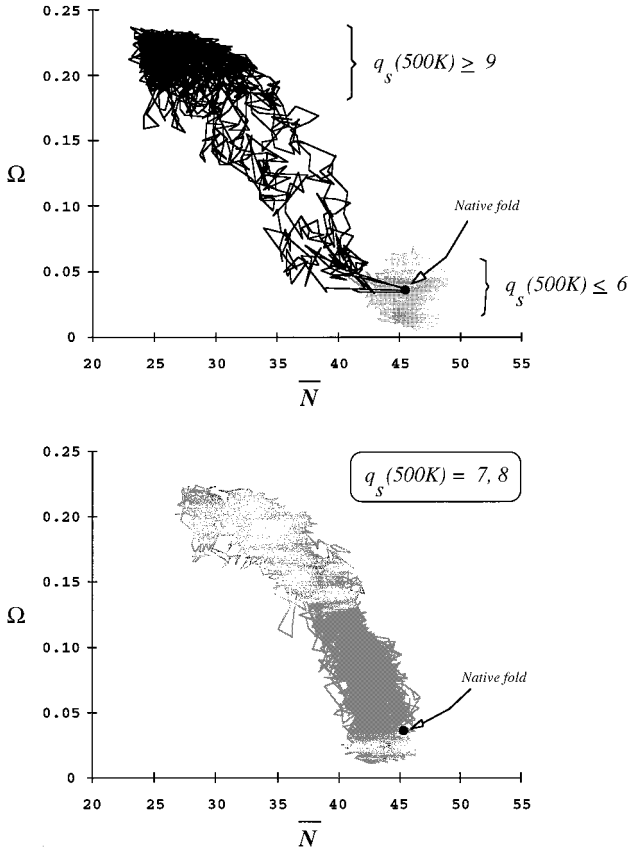

Figure 4. Molecular shape map for the in vacuo dynamics of lysozyme ions at 500 K. The top diagram gives the results for the high-charge states ($q_s \geq 9$) and the low-charge states ($q_s \leq 6$). The bottom diagram gives the results for the intermediate-charge states ($q_s = 7$ and 8). These latter states span the critical charge for unfolding, q_s*, because they show coexistence of both folded (compact) and unfolded (elongated) intermediates.

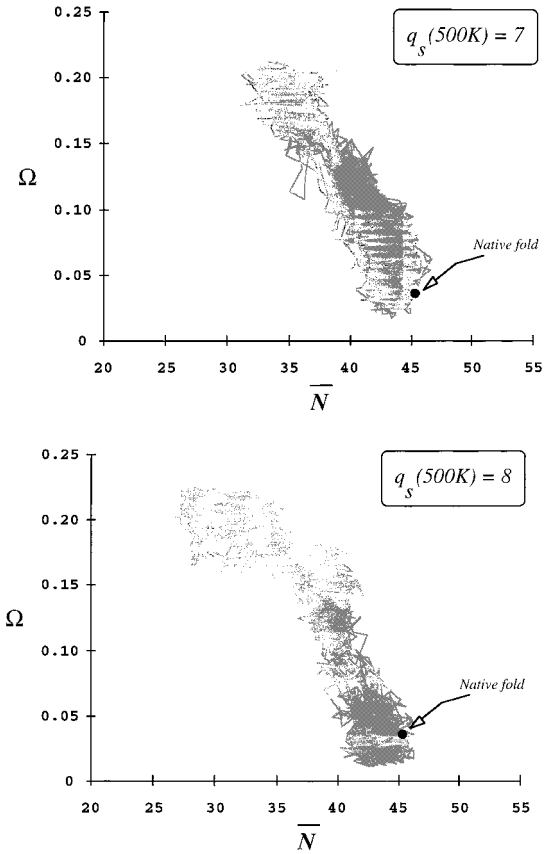

Figure 5. Molecular shape map for the in vacuo dynamics of lysozyme ions with near-critical charge at 500 K. The two diagrams span the map displayed in Figure 4 (bottom). Each diagram superimposes the results for four independent MD runs started with different randomizations of the initial velocities at $T = 500$ K. The results indicate a slightly smaller degree of unfolding at $q_s = 7$ than at $q_s = 8$.

distinguished in terms of self-entanglements. As commented before, the simultaneous use of $\bar{N}$ and Ω provides a more discriminating approach to following changes in molecular shape along the unfolding transition.

The single trajectories in Figures 2 and 3 are not enough to estimate the critical charge needed for unfolding at 500 K. Depending on the initial velocities, unfolding may or may not take place for some q_s values. For this reason, we have monitored multiple trajectories whenever partly unfolded structures were found during the first 3 ns of simulation. Figure 4 displays the ($\bar{N}$,Ω) maps produced with 15 MD runs at 500 K.

Low-charge states correspond to the gray area surrounding the shape of the native fold in the top diagram of Figure 4. This region comprises four trajectories with $q_s = 5$ and 6, where unfolding is never observed. These trajectories fluctuate among conformers with the same degree of compactness and folding complexity as the native fold in the crystal. In this sense, these are *quasinative structures*. In terms of root-mean-square deviations (σ_{rmsd}), these structures can differ from the native fold. However, in this case, values up to $\sigma_{rmsd} = 4$ Å do not necessarily imply the presence of significant differences in fold. In low-charge states, large σ_{rmsd} values are due to local conformational rearrangements that maintain most of the secondary structure and a great deal of the tertiary contacts. Because $\bar{N}$ and Ω are descriptors of large-scale shape features, we can infer the occurrence of quasinative folds despite local distortions from the native configuration.

High-charge states are associated with the black area found at high Ω and low $\bar{N}$ values in Figure 4 (top diagram). This region contains the results of four trajectories with $q_s = 9$ and 11 that undergo fast unfolding (i.e., with only short-lived, partly unfolded transients).

The lower diagram in Figure 4 displays the behavior that characterizes (and identifies) the critical charge for unfolding, q_s*. The diagram contains eight trajectories, with four different MD runs for $q_s = 7$ and 8 each. For each of the charges, we find trajectories that lead to some degree of unfolding. However, complete unfolding never takes place. Instead, we find persistent partly unfolded intermediates, as well as short-lived transients with quasinative folding features (regardless of the corresponding σ_{rmsd} values). The difference is very sharp, because none of these features is observed in the MD trajectories for lysozyme ions with $q_s \geq 9$ and $q_s \leq 6$. As the figure shows, the states $q_s = 7$ and 8 span precisely the "gap" in the ($\bar{N}$,Ω) map left by the low- and high-charge states (cf. Figure 4, top diagram).

The states $q_s = 7$ and 8 span therefore the intermediate regime where folded and unfolded conformers coexist at $T = 500$ K. As shown in Figure 5, the behavior for each of these two charges appears to be similar, yet, is not identical. The conformers populated by lysozyme in the $q_s = 7$ state appear to be slightly less unfolded than those observed for $q_s = 8$. In this sense, $q_s = 8$ covers the complete range of structures stretching from the native fold to almost the fully unfolded conformers in the high-charge regime.

In summary, we can estimate a critical charge for unfolding as q_s*$\sim$7.5 $\pm$ 1.0 at 500 K. This value is in close agreement

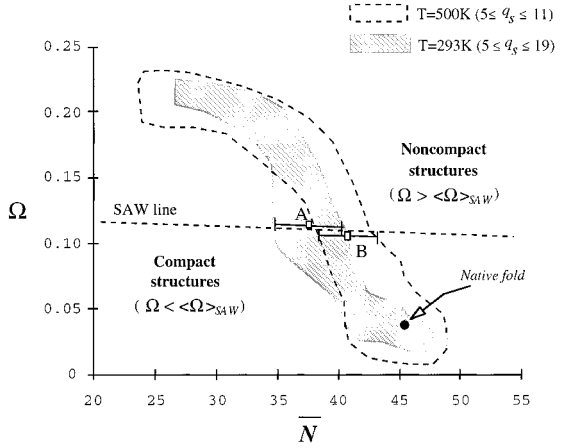

Figure 6. Schematic configurational "phase" diagram for conformers of lysozyme ions at low and high temperature. [Conformers are classified as *compact* or *noncompact* according to their position relative to the "*SAW line*," indicating the behavior of the configurationally averaged descriptors, $\langle\Omega\rangle_{SAW}$ and $\langle\bar{N}\rangle_{SAW}$, for self-avoiding walks.[23] The A and B bars represent the region of $\bar{N}$ values for the critical charges at 293 and 500 K, respectively. Compact conformers found near the native fold represent fluctuations with quasinative folding features, despite root-mean-square deviations of up to 4 Å. Note that higher temperature shifts slightly the shape features associated with the critical charge, but the shapes for conformers with high- and low-charge states remain basically unchanged. This indicates that, over a range of temperatures, the mechanism for unfolding is unaffected. Total ion charge and temperature appear to play reverse roles in the triggering of an unfolding transition.]

with experiments, which indicate $q^* \sim 9$ for in vacuo lysozyme.[3] Because the same model predicts a critical charge of $q_s^* \sim 15 \pm 1$ at $T = 293$ K,[14] it is clear that a realistic simulation must take into account in-flight heating of protein ions. Our results suggest that an effective temperature slightly below 500 K could account for an experimental q^* value between 8 and 9, as inferred from the data in ref 3.

The $(\bar{N},\Omega)$ maps characterize the conformational space accessible to lysozyme ions at 500 K in terms of molecular shape. In this sense, Figures 4 and 5 provide a (single-molecule) configurational "phase" diagram for the unfolding transition. From these diagrams, a pattern emerges in which a number of conformational rearrangements are favored, whereas others are excluded. For example, a sudden melt-down of secondary structure does take place at 500 K while maintaining the globularity of the chain. (Such a process may appear only in the form of very short transients for high-charge states at 500 K.)

The nature of the accessible configurations can be appreciated from Figure 6. Here, we compare the $(\bar{N},\Omega)$ maps for conformers of lysozyme ions at 293 and 500 K. The region enclosed by the dashed curve spans the intermediate structures and persistent transients observed at 500 K for ions with charges $5 \le q_s \le 11$. The gray region encloses the persistent conformers for lysozyme ions with charges $5 \le q_s \le 19$ at room temperature.[5b] Short-lived transients are not included in either area. To assess the relative compactness of a conformer, we have also added the results for self-avoiding walks (SAWs) in the continuum.[34] The configurationally averaged shape descriptors $(\langle\bar{N}\rangle,\langle\Omega\rangle)$ for off-lattice SAWs follow the dashed line in Figure 6. This line represents walks of 129 residues, with a constant bond length of 3.8 Å and a variable excluded-volume interaction. Below the line, we find *compact* configurations, i.e. those produced by a dominant attractive interaction.[23,24] Above the dashed line, polymers are *not compact* because their shapes can only be produced by a dominant repulsive (distance-dependent) interaction.

The position of the "SAW line" in Figure 6 is consistent with our estimation of the critical charge for unfolding q_s^*. The SAW line intersects the region with a fast increase in asphericity and clearly separates the configurations corresponding to fully unfolded lysozymes (high-charge states) from the structures with quasinative folding features (low-charge states). The region assigned to the critical charge q_s^* (cf. Figure 5) represents, indeed, the transition from compact to noncompact conformers.

Figure 6 also indicates that the mean molecular shapes for the conformers in high- and low-charge states are virtually identical at $T = 293$ and 500 K. The only differences lie in the actual values of the charges associated with the species that remain quasinative or unfold at each temperature. In other words, once full unfolding takes place, the conformers reached at 293 and 500 K are similar in terms of molecular shape. (Fluctuations about those conformers are, as expected, larger at 500 K. Note, however, that these fluctuations may cause large σ_{rmsd} values but not necessarily distinct large-scale folding features.) The only significant structural deviations appear in the region for the critical charge. In this case, the degree of chain entanglement for ions with charge q_s^* is slightly increased as the temperature increases.

Final Remarks. We have found that the Coulomb repulsion between charged residues may not necessarily cause unfolding in vacuo. There appears to exist a critical value for the total charge below which protein conformers can persist in tertiary folds not far from the native, even in absence of solvent. In other words, conformations for low-charge ions can be regarded as (possibly large) fluctuations around a robust "attractor" with quasinative folding.

Moreover, we have found that the value for this critical charge decreases with the temperature. This result suggests that the in-flight unfolding of a charged protein can be significantly facilitated by collisional heating with a surrounding gas. The findings are also consistent with the known behavior in some random polyampholytes, which exhibit a critical charge and temperature for the transition from collapsed chains to chains in the self-avoiding walk regime.[35−38]

We have also found that, depending on initial conditions, folded and unfolded conformers can coexist within a range of values of total charge and temperature. These observations suggest that the unfolding of protein ions cannot be represented as a superposition of closely related paths leading to similar conformers but, rather, the transition involves a population distribution over different states.[39] In this sense, the "unfolded" states observed experimentally would include partly folded (e.g., configurationally trapped) structures as well as fully denatured conformations. This notion is in agreement with the results of recent simulations for neutral proteins, where unfolding was monitored over a manifold of trajectories.[40,41]

By using the $(\bar{N},\Omega)$ maps, we have uncovered similarities in the large-scale molecular shape features associated with low- and high-charge states at very various temperatures. This behavior may be related to sequence-specific effects in lysozyme and not found in random polyelectrolytes. The configurational "phase" diagram in Figure 6 strongly suggests that the *same* global mechanism for unfolding lysozyme by Coulombic repulsion is maintained over a range of temperatures. Collisional heating would then appear to *magnify* (or activate) the unfolding pathways already present in charged species at low temperature. This observation recalls a notion akin to the thermodynamic

hypothesis of "corresponding states." According to this analogy, a lysozyme ion with charge $q \sim 16$ would be in an equivalent (conformational) state at $T = 293$ K as an ion with $q \sim 9$ at 500 K. The existence of similar "equivalences" in the conformational spaces of other proteins would be an interesting insight into the nature of the folding−unfolding transition.

Acknowledgment. We thank I. Velázquez for his collaboration during the initial stages of this project and N. D. Grant for comments on the manuscript. G.A.A. acknowledges support from the Natural Sciences and Engineering Research Council of Canada. C.T.R. thanks the TFR (Swedish Technical Research Council), and O.T. is grateful to the NFR (Swedish Natural Sciences Council) for financial support.

References and Notes

(1) Shelimov, K.; Jarrold, M. F. *J. Am. Chem. Soc.* **1996**, *118*, 10313; **1997**, *119*, 2987.

(2) Gross, D. S.; Schnier, P. D.; Rodríguez-Cruz, S. E.; Fagerquist, C. K.; E. K. Williams *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 3143.

(3) Valentine, J. S.; Anderson, J. G.; Ellington, A. D.; Clemmer, D. E. *J. Phys. Chem. B* **1997**, *101*, 3891.

(4) McLafferty, F. W.; Guan, Z.; Haupts, U.; Wood, T. D.; Kelleher, N. L. *J. Am. Chem. Soc.* **1998**, *120*, 4732.

(5) (a) Jarrold, M. F. *Acc. Chem. Res.* **1999**, *32*, 360. (b) Hoaglund-Hyzer, C. S.; Countermann, A. E.; Clemmer, D. E. *Chem. Rev.* **1999**, *99*, 3037. (c) Jarrold, M. F. *Annu. Rev. Phys. Chem.* **2000**, *51*, 179.

(6) Konermann, L. *Sci. Prog.* **1998**, *81*, 123.

(7) 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.

(8) Reimann, C. T.; Sullivan, P. A.; Axelsson, J.; Quist, A. P.; Altman, V.; Roepstorff, P.; Velázquez, I.; Tapia, O. *J. Am. Chem. Soc.* **1998**, *120*, 7608.

(9) Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2426.

(10) Dill, K. A.; Chan, H. S. *Nat. Struct. Biol.* **1997**, *4*, 10.

(11) Dobson, C. M.; Karplus, M. *Curr. Opin. Struct. Biol.* **1999**, *9*, 92.

(12) Przybylski, M.; Glocker, M. O. *Angew. Chem., Int. Ed. Eng.* **1996**, *35*, 806.

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

(14) Reimann, C. T.; Velázquez, I.; Tapia, O. *J. Phys. Chem. B* **1998**, *102*, 9344.

(15) (a) Reimann, C. T.; Velázquez, I.; Bittner, M.; Tapia, O. *Phys. Rev. E* **1999**, *60*, 7277. (b) Arteca, G. A.; Velázquez, I.; Reimann, C. T.; Tapia, O. *Chem. Phys. Lett.* **2000**, *327*, 245.

(16) Mao, Y.; Woenckhaus, J.; Kolafa, J.; Ratner, M. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **1999**, *121*, 2712.

(17) Arteca, G. A. *Biopolymers* **1995**, *35*, 393; *Macromolecules* **1996**, *29*, 7594.

(18) Baumgärtner, A. *J. Chem. Phys.* **1993**, *98*, 7496.

(19) Stasiak, A.; Katritch, V.; Bednar, J.; Michoud, D.; Dubochet, J. *Nature* **1996**, *384*, 122.

(20) Katritch, V.; Bednar, J.; Michoud, D.; Scharein, R. G.; Dubochet, J.; Stasiak, A *Nature* **1996**, *384*, 142.

(21) Vologodskii, A. V.; Crisona, N. J.; Laurie, B.; Pieranski, P.; Katritch, V.; Dubochet, J.; Stasiak, A. *J. Mol. Biol.* **1998**, *278*, 1.

(22) Arteca, G. A. *Biopolymers* **1993**, *33*, 1829; **1996**, *39*, 671.

(23) Arteca, G. A.; Velázquez, I.; Reimann, C. T.; Tapia, O. *Phys. Rev. E* **1999**, *59*, 5981.

(24) Arteca, G. A.; Velázquez, I.; Reimann, C. T.; Tapia, O. *J. Chem. Phys.* **1999**, *111*, 4774.

(25) Reimann, C. T.; Velázquez, I.; Tapia, O. *J. Phys. Chem. B* **1998**, *102*, 2277.

(26) Velázquez, I.; Reimann, C. T.; Tapia, O. *J. Am. Chem. Soc.* **1999**, *121*, 11468.

(27) Marchi, M.; Ballone, P. *J. Chem. Phys.* **1999**, *110*, 3697.

(28) (a) Åqvist, J.; van Gunsteren, W. F.; Leijonmarck, M.; Tapia, O. *J. Mol. Biol*. **1985**, *183*, 461. (b) van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulation—GROMOS Library Manual*; Biomos: Groningen, The Netherlands, 1987.

(29) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(30) Miteva, M.; Demirev, P. A.; Karshikoff, A. D. *J. Phys. Chem. B* **1997**, *101*, 9645.

(31) (a) Arteca, G. A.; Caughill, D. I. *Can. J. Chem*. **1998**, *76*, 1402. (b) Arteca, G. A. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 550.

(32) Arteca, G. A.; Tapia, O. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 642.

(33) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graph.* **1996**, *14*, 51.

(34) Arteca, G. A. *Phys. Rev. E* **1994**, *49*, 2417.

(35) Kantor, Y.; Kardar, M. *Europhys. Lett.* **1994**, *27*, 643; **1994**, *28*, 169.

(36) Kantor, Y.; Kardar, M.; Ertas, D. *Physica A* **1998**, *249*, 301.

(37) Wolfling, S.; Kantor, Y. *Phys. Rev. E* **1998**, *57*, 5719.

(38) Dobrynin, A. V.; Rubinstein, M.; Obukhov, S. P. *Macromolecules* **1996**, *29*, 2974.

(39) Konermann, L.; Douglas, D. J. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 435.

(40) Lazaridis, T.; Karplus, M. *Science* **1997**, *278*, 1928.

(41) Kazmirski, S. L.; Daggett, V. *J. Mol. Biol.* **1998**, *277*, 487.