Protein denaturation in vacuo: Mechanism for centrifugal unfolding of neutral lysozyme

ARTICLE *in* THE JOURNAL OF CHEMICAL PHYSICS · DECEMBER 2001

Impact Factor: 2.95 · DOI: 10.1063/1.1412865

2 AUTHORS, INCLUDING:

Orlando Tapia
Uppsala University
236 PUBLICATIONS **4,270** CITATIONS

SEE PROFILE

Protein denaturation in vacuo: Mechanism for centrifugal unfolding of neutral lysozyme

Gustavo A. Arteca and O. Tapia

Citation: J. Chem. Phys. **115**, 10557 (2001); doi: 10.1063/1.1412865
View online: http://dx.doi.org/10.1063/1.1412865
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v115/i22
Published by the American Institute of Physics.

Protein denaturation *in vacuo*: Mechanism for centrifugal unfolding of neutral lysozyme

Gustavo A. Arteca[a)]
Département de Chimie et Biochimie, Laurentian University, Ramsey Lake Road, Sudbury, Ontario P3E 2C6, Canada

O. Tapia
Department of Physical Chemistry, Uppsala University, Box 532, Uppsala S-751 21, Sweden

(Received 4 June 2001; accepted 29 August 2001)

Experiments on anhydrous proteins diffusing in a low-pressure gas reveal transitions that share common features with the unfolding and refolding processes observed in solution. These phenomena force us to re-evaluate the specific role played by solvent on large-scale protein rearrangements. Computer simulations, in combination with tools for molecular shape analysis, provide insights into *in vacuo* processes. In this work, we deal with one particular aspect of this problem: the conditions and the mechanism for spontaneous unfolding in a globally neutral protein. When coupling the protein to a simulated thermal bath, the flow of energy between rotational modes and internal modes can produce a centrifugal effect leading to unfolding. Using hen lysozyme as a system, we study the reproducibility of the unfolding transition and its dependence on the bath relaxation constant. In addition, we monitor the evolution of large-scale molecular shape features (e.g., chain entanglements) that take place during unfolding. Our results show that a change in bath relaxation constant affects the time scale of the transition and the population of intermediates, without changing the basic unfolding mechanism. We also discuss possible cycles of unfolding–refolding transitions, and their implications for our understanding of the denatured state. © *2001 American Institute of Physics.* [DOI: 10.1063/1.1412865]

I. INTRODUCTION

Intact anhydrous proteins (or protein ions) can be transferred to the gas phase by using laser desorption and electrospray ionization.[1–10] Measurements of ion-drift mobility[1,3,8–10] and surface imprinting[11,12] provide, in turn, tools for their low-resolution structural characterization. These experiments show that conformers with nativelike characteristics are possible (at least, metastable) *in vacuo*. Results also indicate that refolding into compact structures is possible even without water, a notion debated since 1995.[13] Using computer simulations mimicking gas-phase migration, we can gain insights into the principles that govern many of these processes. In this work, we address one of them: the unfolding transition in a globally neutral, anhydrous protein.

The proper simulation of proteins in gas-phase migration experiments requires protocols with *in vacuo boundary conditions*. Under these conditions, biomolecules can rotate and translate freely. Thermalization by collisions with a low-pressure buffer gas can then be incorporated through a global coupling to a simulated thermal bath. Many factors may trigger unfolding. In sufficiently charged (i.e., protonated) protein ions, unfolding proceeds by Coulombic repulsion in conditions that simulate low *p*H denaturation.[14] A more uncontrolled transition can be caused by allowing extreme temperatures.[15–18] In this work, we deal with conditions that promote a smoother room temperature unfolding of a neutral protein. Previous approaches have studied the effect of an external bias.[19,20] This technique can mimic experiments with single-molecule manipulation,[21–25] or the situation of confined proteins,[26,27] but it does not fit the conditions prevalent in gas phase. In this case, ''spontaneous'' unfolding is possible by a proper manipulation of the relaxation constant coupling the protein to the thermal bath. This procedure, referred to as ''centrifugal unfolding,'' [28] has been used to generate unfolded conformers of hen lysozyme for studies of refolding relaxation.[27–31] Although a qualitative pattern of relaxation from unfolded structures begins to emerge,[30–33] much less is known on the unfolding process itself. In this work, we systematically study this effect. Our first goal is to analyze the reproducibility of the unfolding transition with *in vacuo* boundary conditions, and the role played by the relaxation constant. In addition, we discuss the pattern of molecular shape changes during centrifugal unfolding using descriptors of chain entanglement. Our analysis provides insight into some fundamental characteristics of the denatured state, and also for understanding the differences between unfolding and refolding pathways.[34,35]

II. COMPUTER SIMULATION OF UNFOLDING AND CHARACTERIZATION OF UNFOLDED CONFORMERS

The immediate goals of our study are: (i) to establish the conditions that control the onset of unfolding in a well-defined molecular dynamics (MD) protocol for neutral proteins, and (ii) to characterize the large-scale structural features of the resulting unfolded continuum. The initial

a)Author to whom correspondence should be addressed. Electronic mail: gustavo@nickel.laurentian.ca

motivation for this work is a recent result: under specific simulation conditions, anhydrous lysozyme is known to unfold rapidly at room temperature without the help of an external force.[28] It is unclear at this moment whether this procedure, referred to as ''centrifugal unfolding,'' is akin to a realistic denaturing process *in vacuo*. Nevertheless, the simulation results agree with some available evidence: (i) The mean lengths and cross sections of the unfolded conformers compare well with experimental values,[12] and (ii) the transition proceeds stepwise within a much shorter time scale than the one observed in solution,[36] as would be expected for gas phase.[8–10] So far, we have no detailed understanding of the mechanism at play in this simulation. Our present objective is pragmatic: to establish whether or not centrifugal unfolding is a reproducible and reliable approach to generate a well-defined ensemble of denatured conformations. Below, we discuss the simulation conditions and the tools used for the characterization of large-scale shape features.

A. Molecular dynamics simulations

We use constant-temperature MD trajectories for anhydrous lysozyme. All simulations start from the crystal structure of hen lysozyme (PDB code *1 hel*), a single chain protein with 129 residues. During the simulations, we maintain the chain connectivity, including all disulfide bridges. The protein is modeled with the Gromos 87 force field,[37,38] which includes explicit polar hydrogens and electrostatic hydrogen bonding by way of the 37D4 parameter set. The potential energy parameters for the protein do not depend explictly on the environment, hence they can be used for *in vacuo* simulations. The Gromos force field includes standard contributions:

$$V = V_{\text{bond}} + V_{\text{bend}} + V_{\text{imp-dih}} + V_{\text{tor}} + V_{\text{Lj}} + V_{\text{elec}}, \tag{1}$$

where V_{bond}, V_{bend}, and $V_{\text{imp-dih}}$ are the harmonic potentials for bond vibrations, bendings, and improper dihedral motions, respectively; the V_{tor} term corresponds to dihedral torsions. The last two terms in Eq. (1) correspond to the nonbonded pair potential (modeled by a Lennard-Jones 6-12 function), and the electrostatic interaction between point charges. Note that the force field includes amino acid charge distributions, yet the protein is globally neutral.[38]

We are interested in the factors that control the onset of unfolding as an *early response* behavior under the simulation conditions. Our conclusions are based on an ensemble of 2-ns long MD trajectories; though the trajectories are short for unfolding phenomena in solution, they are appropriate for the present goals *in vacuo*. The protocol employs an integration step of $\Delta t = 2$ fs in the Verlet algorithm. Lennard-Jones interactions include a 13 Å distance cutoff. Independent trajectories are generated by changing the set of initial random velocities at a bath temperature $T_0 = 293\,\text{K}$. In order to mimic the situation *in vacuo*, we do not use periodic boundary conditions, nor do we remove inertial rotations and center-of-mass translations.

The key control parameter for the present analysis is the relaxation constant τ that regulates the coupling to the bath. To compare with previous simulations, we use a Berendsen thermostat,[39] where the kinetic (instantaneous) temperature follows Newton's cooling law: $\mathrm{d}T/\mathrm{dt} = (T_0 - T)/\tau$, with T_0 the bath temperature. Large τ values ensure slow cooling, i.e., a *weak* coupling to the bath. This coupling modifies the instantaneous velocities during the MD run with a scaling factor λ at every time step of length Δt:[39]

$$\lambda = \left[1 + \frac{\Delta t}{\tau} \left(\frac{T_0}{T} - 1 \right) \right]^{1/2}. \tag{2}$$

The shortcomings of this approach are well known:[40,41] it does not produce a canonical (NVT_0) ensemble distribution. This is not a concern for the present work, since we do not attempt to map all unfolded conformations or evaluate canonical averages. Instead, we deal with a situation in which the method *is* useful, i.e., a tool for changing the state of a macromolecule by means of a smooth dissipation mechanism. This model couples the macromolecule globally to the heat bath while causing the minimum disturbance to its local dynamics.

The trajectories resulting from the above algorithm allow energy transfer from internal degrees of freedom to collective vibro-rotations. This is a key feature that makes this approach useful as a practical tool to elicit unfolding in a neutral protein. The principle at work is as follows.[28,42] Within the force field, Eq. (1), bending modes are effectively coupled through the dihedral torsions. Moreover, rotations can persist over long periods when using *in vacuo* boundaries conditions, and the magnitude of these rotations is affected by the bath relaxation. As a result, nonrigid rotations about the inertial axes can affect over time internal degrees of freedom, in particular those involving large-scale torsions and bending.[43] The end result is a nonrigid molecular rotor *in vacuo*, one where energy flows between localized and globally extended vibrational modes depending on the value of τ. Under these conditions, a protein can eventually be subject to spontaneous (and sustained) spinning about an inertial axis. This effect results in weak, yet systematic, ''centrifugal forces'' acting persistently on the same type of collective modes. A transition occurs if a large-scale vibrational mode for unfolding is among those modes affected by the energy transfer. A similar phenomenon is known to take place in smaller molecules, where bonds can be broken by forced rotations using what effectively is an ''optical centrifuge.'' [44–46] Whether or not this phenomenon is observable in a macromolecule within a 2-ns MD window will depend on the particular protein structure, and on the degree of ''rotational activation'' of an accessible unfolding mode. (Note that unfolding half-lives shorter than 10 ns are indeed possible for small α-helical proteins in solution.[47]) As we show below, lysozyme may have favorable features that make a fast centrifugal unfolding possible. By regulating the rotational activation with the value of the bath coupling τ, it is possible to denature this neutral protein at room temperature within a short time scale. Given the intrinsic limitations of the method, we base our conclusions on the pattern emerging from many independent trajectories.

B. Large-scale molecular shape features

Conformational snapshots extracted during the MD trajectories must be analyzed in order to monitor the onset of unfolding. We focus on properties that discriminate among chains with different folds, yet have similar molecular sizes or degree of compactness. We avoid root-mean-square deviations ($rmsd$), mean sizes, and the content of secondary structure. Instead, we use descriptors that measure the absolute and relative complexity of chain self-entanglements.

A *self-entangled* chain exhibits projected bond-bond crossings (''overcrossings'').[48–51] Overcrossings are a powerful tool to study nontrivial topologies (e.g., knotted DNA),[52,53] but they are also useful as geometrical descriptors for linear chains with transient entanglements. The degree of self-entanglement, expressed in terms of the distribution of overcrossings,[50,51] conveys essential aspects of the three-dimensional shape of a single chain, whether knotted or not.

For a given conformation K, we compute numerically the probability distribution $\{A_N^{(K)}\}$ for $N \geqslant 0$ overcrossings, averaged over a large number of random directions.[50] By construction, descriptors based on $\{A_N^{(K)}\}$ incorporate both the geometry and connectivity of chain; they are also rotationally and translationally invariant. As an absolute descriptor of entanglement, we use the first moment of the distribution, or mean overcrossing number, $\overline{N}$. From the operational viewpoint, this ''overcrossing index'' can be conceived as follows. Consider a polymer chain at an arbitrary internal configuration, corresponding to the equivalence class of rigid translations and rotations. Imagine that we now make a projection of the main backbone of the polymer along a generic direction. For simplicity, this plane can be thought as tangent to the smallest sphere that contains the backbone. Then, let us compute the number of ''crossings,'' N, between pairs of backbone bonds at this projection. Finally, the mean overcrossing number $\overline{N}$ is computed as an average of N the above number of projected crossings over all possible projections. Computationally, the averaging can be achieved by scanning over a large number of projections randomly distributed on the enclosing sphere mentioned before.[50] Also, the $\overline{N}$ value can also be computed directly using a path-integral representation over all pairs of overcrossing segments or ''bonds.'' [54] As a *relative descriptor* of entanglement, we use the root-mean-square deviation of overcrossing probabilities between the instantaneous conformer K and a reference one K^*:[55]

$$\sigma_A(K,K^*) = \left[\sum_{N=0}^{N_{\max}} (A_N^{(K)} - A_N^{(K^*)})^2 \right]^{1/2}. \tag{2}$$

In our particular case, we take K^* as the native state in the crystal. As a chain unfolds, we expect $\overline{N}$ to decrease depending on the extent of residual secondary structure. (Note that $\overline{N}=0$ in a rod.) Similarly, the relative descriptor σ will increase from typical values corresponding to the statistical fluctuations in a native state and its mutations, e.g., $\sigma(K,K^*) \approx 0.035 \pm 0.015$.[55]

As a complementary tool, one can also use a descriptor of chain compactness that is *a priori* uncorrelated with $\overline{N}$. A useful parameter is the chain *asphericity* Ω, given in terms of the principal moments of inertia $\{I_i\}$ as:[56] $\Omega = \{(I_1-I_2)^2 + (I_1-I_3)^2 + (I_3-I_2)^2\}/2(I_1+I_2+I_3)^2$. As a chain changes from spheroidal in the native state to prolate in the unfolded state, Ω changes from $\Omega=0$ to $\Omega \approx 1/4$. The pair of descriptors $(\overline{N},\Omega)$ defines a simple and convenient shape space for the analysis of unfolding–refolding transitions.[30–33]

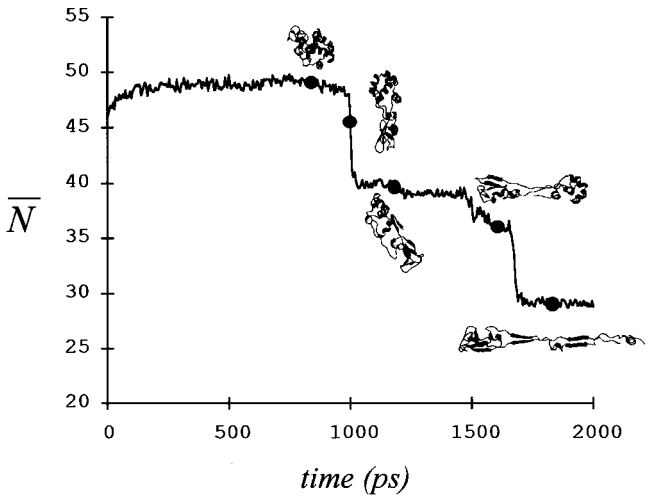

FIG. 1. Changes in chain entanglement (measured by the mean overcrossing number) in a MD trajectory for the *in vacuo* unfolding of neutral lysozyme at $T=293$ K. The run has a bath coupling constant $\tau=40$ fs. The selected snapshots illustrate the two-step formation of elongated conformers. Unfolding is due to the effect of ''centrifugal'' forces related to molecular rotations.

III. EFFECT OF BATH RELAXATION ON THE CENTRIFUGAL UNFOLDING OF NEUTRAL *1 hel*-LYSOZYME

Using the protocol in Sec. II, the literature reports unfolding in a 1.2-ns trajectory with $\tau=10$ fs, and no unfolding for $\tau=100$ fs.[28] In the latter case, the lysozyme fold rearranges itself in a slightly more compressed form compared to the crystalline native state. We refer to this metastable ensemble of conformers as the *in vacuo* native structure (IVNS).[28–30] In this section, we study systematically the stability of the IVNS over ensembles of MD trajectories when the bath relaxation varies from $\tau=5$ fs (a strong coupling) to $\tau=100$ fs (a weak coupling). For each τ value, we have collected at least ten stable trajectories, where temperatures remain at $T=292\pm1$ K during 2 ns. For proper comparison, we use the same seeds for the random number generator that initializes the velocities in all trajectories.

As an illustration, let us discuss briefly a typical result. Figure 1 shows the evolution of the mean overcrossing number during a 2-ns MD trajectory with $\tau=40$ fs. Selected snapshots are displayed at times indicated by the black dots. (Snapshots have been generated with the program *Molmol*.[57]) During the initial 50 ps, the structure gains slightly in entanglement due to the electrostriction caused by the lack of solvent screening *in vacuo*.[10,28] Thereafter, lysozyme remains in the IVNS until a sharp transition at $t \approx 1000$ ps. As indicated by the snapshots, at this point the conformation is changed from a rather compact globule into an elongated object with little change in its local secondary structure. The

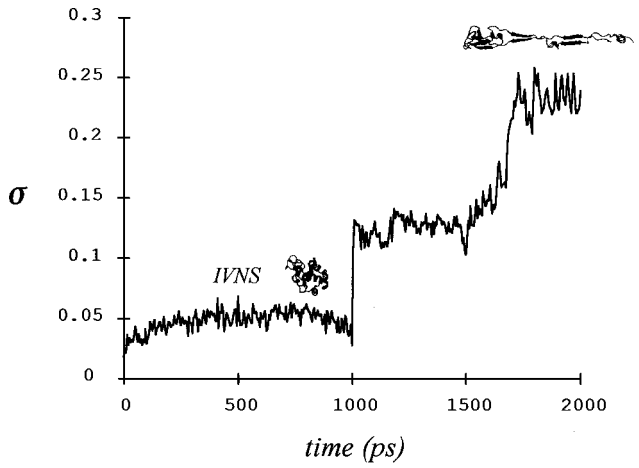

FIG. 2. Changes in relative entanglement, measured by the root-mean-square deviation (*rmsd*) in overcrossing probabilities with respect to the native state of lysozyme in the crystal. The term IVNS stands for "*in vacuo* native structure," which is a compressed (and metastable) version of the initial native state. Note the reduction in *rmsd* just prior to unfolding. The results correspond to the same simulation as in Fig. 1.

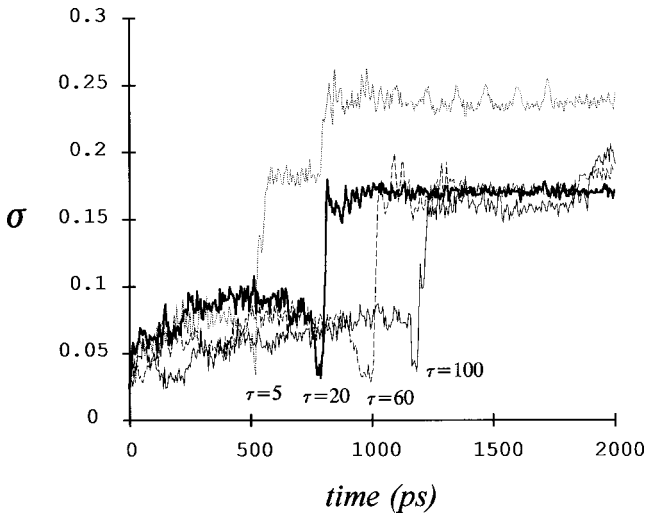

FIG. 3. Dynamics of relative entanglement for selected unfolding trajectories using various τ values (in fs). In all cases the minimum in σ identifies a maximum similarity to the native state in the crystal prior to the initial unfolding transition. The transition is generally delayed when the τ increases (see text).

motion responsible for this change is a large-scale opening of the lysozyme "hinge" that separates the α- and β-rich domains.[28] This motion is brought about by the onset of a systematic rotation about an inertial axis perpendicular to the approximate symmetry axis of the conformer. (The energy associated with these global rotations can become as large as a third of the total kinetic energy.[28]) These partly unfolded conformers persist until $t \approx 1500$ ps, where a second transition takes place. The fully unfolded conformers at the end of the trajectory exhibit mean lengths comparable to those observed experimentally.[12] At this point, internal rearrangements stop and the only observable motion is a high-angular momentum rotation coupled to a global bow-string vibration. The very elongated shape of the conformers, and their residual β-sheet structure, are maintained by the presence of the disulfide bridges.

The behavior for $\overline{N}$ can be contrasted with that of the relative root-mean-square deviation in overcrossing probabilities, σ, in Fig. 2. The two-step unfolding transition can be recognized in the rapid change of σ near $t \approx 1000$ ps and 1600 ps. This descriptor magnifies the effect of fluctuations in molecular shape; this is particularly evident in the σ-oscillations after full unfolding ($t > 1600$ ps), associated with bow-string vibrations. Figure 2 also indicates that there is a drift in σ during the formation of the IVNS, reaching a maximum of $\sigma \approx 0.05$ at ca. 600 ps. However, σ *decreases* just prior to the first unfolding transition. This behavior is observed throughout the simulations, particularly if $\sigma \geqslant 0.05$ during the IVNS phase. Figure 3 illustrates the trends with some remarkable cases corresponding to MD trajectories computed with relaxation constants $\tau = 5$, 20, 60, and 100 fs. In all cases, the unfolding transition is preceeded by a sharp decrease in the *rmsd* parameter σ. Given that the value at the minimum ($\sigma \approx 0.035$) is typical of crossed *rmsd* for stable mutations of the native state in the crystal,[55] we conclude that *lysozyme adopts conformations with maximal similarity to the initial structure just before unfolding*.

The behavior demonstrated in Fig. 3 suggests some interesting aspects of the unfolding mechanism. First, it is clear that the protein must "undo" the compactization of the IVNS in order to unfold. Our results point to the fact that the accessible paths allowing this decompactization take the protein back to structures resembling the crystalline native state. There appear to be no pathways linking a compactified IVNS directly to the unfolding continuum. From this picture, the notion of native state that emerges is one where the latter structure can be described as *marginally* stable, or at the "threshold" of unfolding.

The results in Fig. 3 also demonstrate another trend: unfolding can occur for high and low τ values, although the sharp transition appears to shift to later times as τ increases. The results over many MD trajectories, depicted in Figs. 4–6, support this view. For the sake of brevity, we show the change in absolute and relative overcrossings for $\tau = 5$ fs (Fig. 4), 10 fs (Fig. 5), and 40 fs (Fig. 6), resulting from the analysis of snapshots extracted every 5 ps. The trajectories for various τ values can be compared directly since we use the same sets of initial velocities. We can summarize our main observations as follows:

(a) When using a very strong coupling ($\tau = 5$ fs, Fig. 4), all stable trajectories show unfolding. The transition appears as a very sharp change in overcrossing number at a time t^*. Note that $\overline{N}$ and σ span a narrow range of values after the transition. Thus, the transition involves only two "states," the IVNS before t^* (characterized by $\overline{N} \approx 49$) and the unfolded state after t^* (characterized by $\overline{N} \approx 28$). The phenomenon of retracing the initial native state before unfolding can be observed clearly in σ, although not in $\overline{N}$. In addition, the onset of unfolding shows a strong synchrony throughout the MD runs, given the small dispersion at $t^* \approx 600 \pm 100$ ps. (A preliminary analysis of this figure was presented in a symposium.[42]) Since this τ value is not far from the integration step ($\Delta t = 2$ fs), these results suggest that a centrifugal

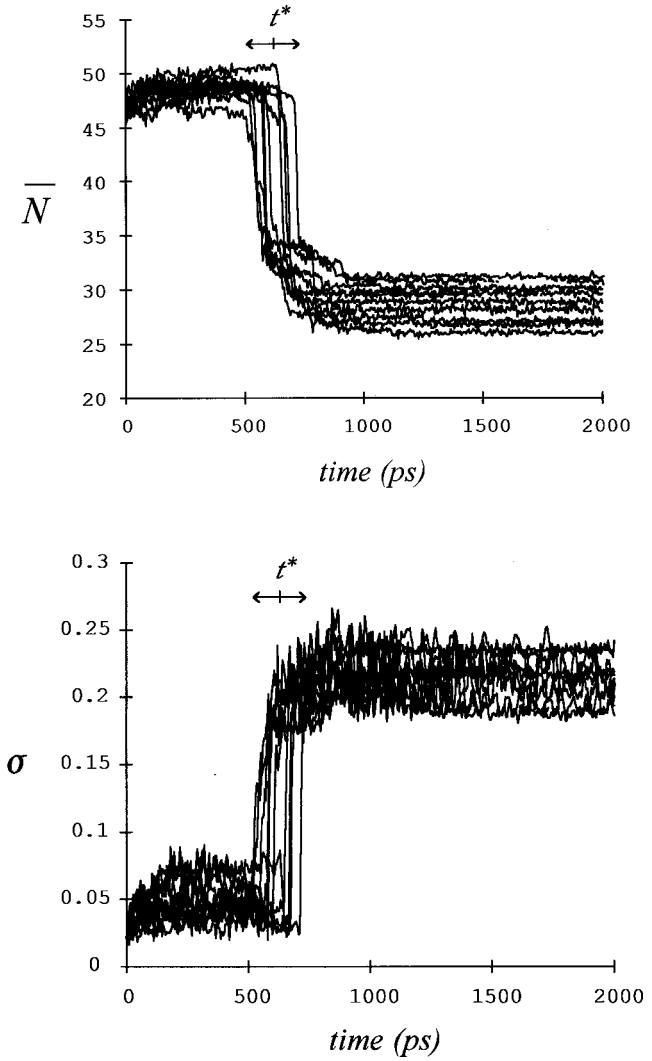

FIG. 4. Results of ten independent MD trajectories for centrifugal unfolding of lysozyme using $\tau=5$ fs. Each trajectory uses a different set of randomized initial velocities at $T=293$ K. There is a sharp unfolding transition at $t^*\sim 600$ ps in both the mean overcrossing number (top) and the *rmsd* in overcrossing probabilities (bottom).

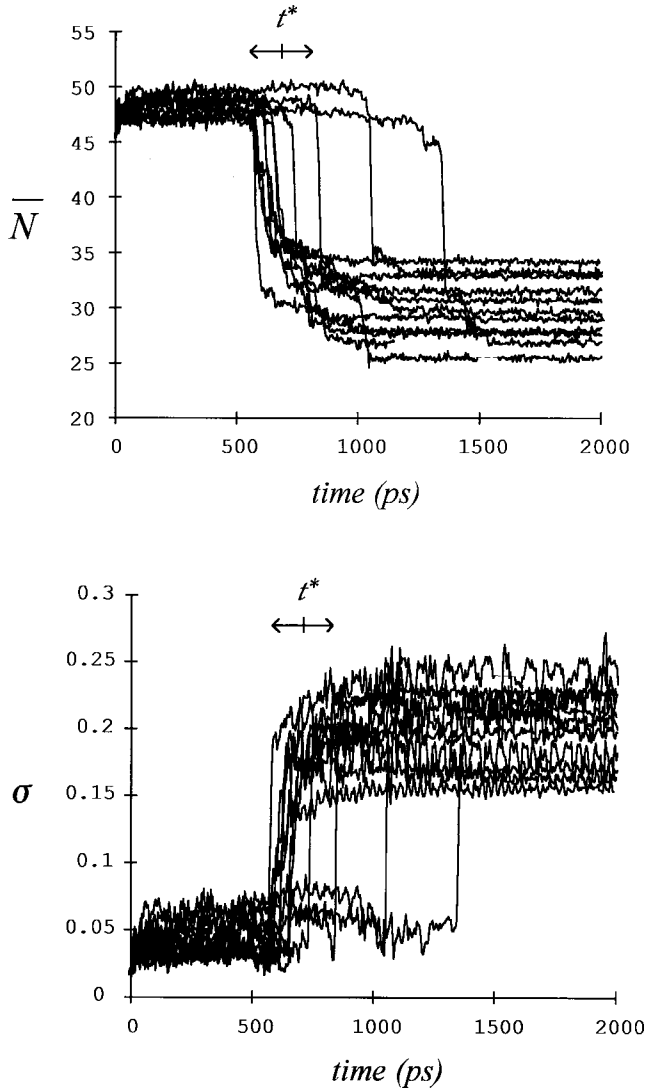

FIG. 5. Results of an ensemble of independent MD trajectories for centrifugal unfolding of lysozyme using $\tau=10$ fs. The onset of the unfolding transition occurs over a larger range of values compared to Fig. 4. Note also the wider range of $\overline{N}$ and σ values after unfolding.

unfolding transition could not possibly occur before ca. 500 ps.

(b) For the next bath coupling ($\tau=10$ fs, Fig. 5), unfolding also occurs for all stable trajectories but with a larger dispersion for the transition time t^*. Roughly half of the MD runs unfold at the time observed for $\tau=5$ fs; in other cases, the transition is delayed up to 1300 ps. Note also that the range of $\overline{N}$ and σ values associated with the unfolded "state" is larger than in the previous case. The slightly larger value of the mean overcrossing number for the unfolded structures ($\overline{N}\approx 30$) suggests the persistence of a larger residual secondary structure.

(c) The first evidence that unfolding can be delayed beyond the present 2-ns window appears when using $\tau=20$ fs. Still, unfolding is the most probable outcome, although it can occur over a larger range of t^* values. (For this reason, we cannot estimate the mean value for t^*.) The transition still appears as a sharp two-state process, with the same shape descriptor values as those in (b). This behavior dominates when using weaker coupling values. Figure 6, corresponding to $\tau=40$ fs, shows that the majority of the trajectories lead to unfolding, but the transition is delayed up to $t^*\approx 1900$ ps. Early transitions do not lead to fully unfolded structures, but rather to partly unfolded conformers characterized by $\overline{N}\approx 38$ and $\sigma\approx 0.15$.

(d) The dynamical behavior for $\tau=60$ fs, 80 fs, and 100 ps follows the pattern observed in (c). With the weakest coupling, lysozyme remains in the IVNS form up to ca. 1100 ps for all trajectories. In addition, only less than half of the trajectories unfold before 2000 ps. These features put the tentative findings in Ref. 28 on firmer ground by providing a quantitative estimation of the stability of the native state under conditions of a weak bath coupling. Finally, note that the range of accessible molecular shapes at the unfolded "state" is wider when using the weakest coupling.

(e) In closing, Fig. 7 collects the unfolding statistics over the entire ensemble of trajectories. The figure shows the mean overcrossing number corresponding to the last configu-

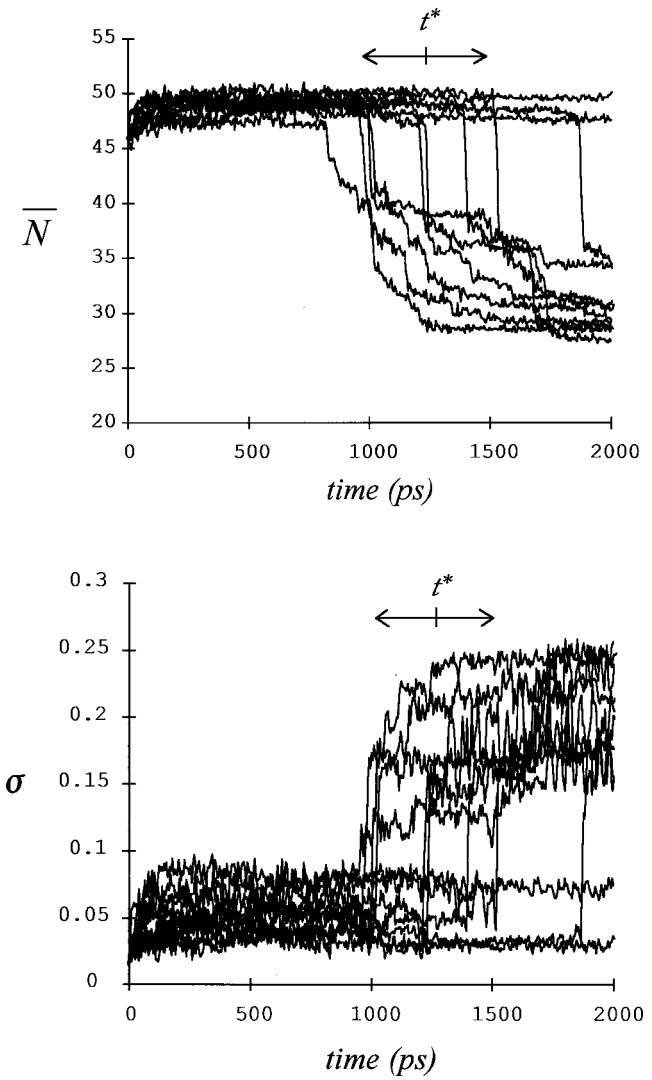

FIG. 6. Results of an ensemble of independent MD trajectories for centrifugal unfolding of lysozyme using $\tau = 40$ fs. The trajectories include the one highlighted in Figs. 1 and 2. Note the IVNS is metastable up to ca. 1000 ps; two trajectories do not lead to unfolding.

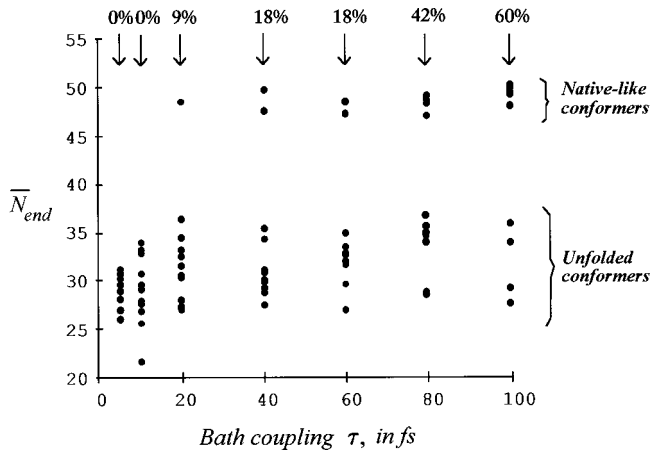

FIG. 7. Unfolding statistics for all MD trajectories and τ values. The figure shows the mean overcrossing number, $\overline{N}_{\text{end}}$, corresponding to the last conformational snapshot of each trajectory. The numbers on top indicate the approximate percentage of trajectories that remain in *folded* states up to 2 ns. Despite differences in the number of unfolding events, the sets of folded and unfolded conformers collected for all τ values exhibit similar molecular shape features.

rational snapshot for each MD run and τ value. On top, we indicate the approximate percentage of trajectories where the nativelike fold is conserved during 2 ns. As the figure show, unfolding becomes a less frequent outcome as τ increases. This result is consistent with the notion that centrifugal unfolding involves energy transfer to large-scale vibrations facilitated by rotational activation. A weaker coupling to the thermal bath reduces the possibility for inducing fast spinning in a protein.

In summary, our main findings in this section are threefold: (1) unfolding is possible for all values of the bath coupling constant; (2) for larger τ values, the unfolding transition is delayed and the population of persistent partly unfolded intermediates increases; (3) the minimum time required for the accumulated effect of centrifugal forces to trigger unfolding is estimated at $\min t^* \sim 500$ ps.

The latter estimate for $\min t^*$ would appear to rule out the possibility that a similar mechanism of centrifugal unfolding could work in solution, where constant solvent collisions should inhibit a long-lived systematic rotation about any given inertial axis. Nevertheless, our main observation is not affected: The net result of weak, yet systematic, forces on lysozyme is to produce a two-state unfolding transition. In this particular work, disruptive forces are simulated by global ''centrifugal'' effects. However, regardless of whether or not such a process can survive long enough in condensed phase, our analysis suggests an *intrinsic structural response of the native fold of lysozyme under the presence of denaturing conditions.*

IV. QUALITATIVE UNFOLDING PATTERN

The results in Fig. 7 suggest that the all unfolded states detected during the simulations share similar molecular shapes despite variations in τ. An equivalent similarity is also apparent among conformers classified as folded states. We can test these conjectures by monitoring the behavior of a different shape descriptor. To this end, we have studied the interrelation between chain entanglements and chain asphericity. By using the $(\overline{N},\Omega)$-map, we can express the protein folding dynamics in terms of two important concepts: the degree of compactness conveyed by Ω, and the organization of tertiary structure conveyed (to some extent) by $\overline{N}$.

The two diagrams in Fig. 8 collect the data for strong couplings ($\tau = 5$ and 10 fs, top) and weaker couplings ($\tau = 80$ and 100 fs, bottom). The thin lines represent the MD trajectories, starting from the native state (black dot). In order to assess the degree of compactness in an unbiased manner, we have added into the diagram the so-called ''SAW line.'' [30,58] The line joins the values for the configurationally averaged descriptors, denoted by $\langle\overline{N}\rangle$ and $\langle\Omega\rangle$, computed in a model of self-avoiding walks (SAW) with variable excluded-volume interaction and the chain length of a lysozyme backbone.[30] This model includes only repulsion; as a result, structures found below the SAW line are compact, since they

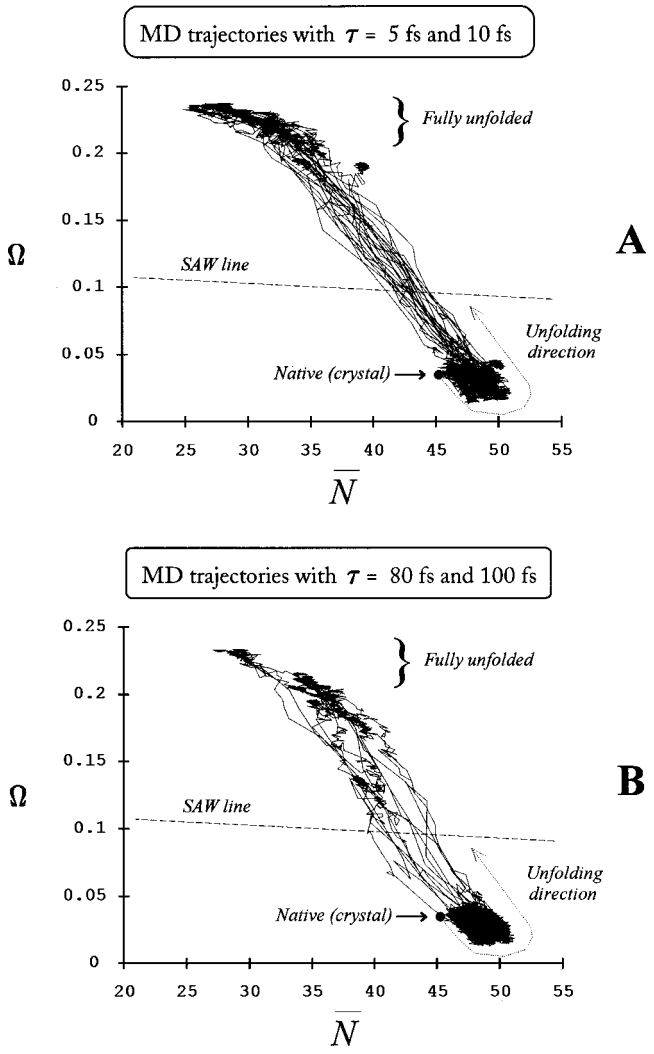

FIG. 8. Pattern of unfolding in an asphericity-and-overcrossing map for trajectories with various τ values. Diagram A shows all trajectories for the stronger couplings ($\tau=5$ and 10 fs). Diagram B displays the trajectories for the weaker couplings ($\tau=80$ and 100 fs). Structures above the SAW line are noncompact (see text). The unfolding process proceeds by a correlated loss of chain compactness (Ω) and entanglements ($\overline{N}$). Changes in bath relaxation affect the persistence of intermediate conformations; however, the essential aspects of the unfolding mechanism appear to depend little on τ.

can only be reached by including an attractive interaction. Structures above the line correspond to noncompact conformers.

In the case of strong coupling (Fig. 8A), the MD runs appear as lines with little noise between $\Omega\approx0.05$ and $\Omega\approx0.15$. These trajectories lead rapidly to the unfolded continuum with few conformers of intermediate compactness. In contrast, the MD trajectories for weak coupling (Fig. 8B) exhibit a distinct population of longer-lived conformers over the SAW line. (The population is small due to the fact that the delayed t^* values make unfolding a less frequent outcome.) Yet, despite differences in persistence, all trajectories cover a similar swath on $(\overline{N},\Omega)$-space. In other words, the conformers' *shape* and the overall *pattern of unfolding* is conserved over changes in τ. The intermediates found when using a weak coupling share the same molecular shape features as the transients found when using a strong coupling. These findings agree with recent observations made in lattice models for unfolding. These simulations indicate that unfolding at low and high temperature proceeds through essentially the same pathways, although their population of intermediates are different.[35]

The common pattern on shape space suggests that the same well-defined process of centrifugal unfolding is at play for all bath relaxations. As indicated by Fig. 8, compactness and chain entanglement are strongly correlated along the unfolding paths. The resulting unfolding mechanism must thus involve a synchronous loss of compactness and tertiary structure. It is worth noting that the emerging process resembles, in a reversed fashion, the conjectured ''optimal'' paths for protein folding.[59,60] As well, our findings fit well with recent simulations in solution, that indicate unfolding pathways dominated by an organized loss of tertiary contacts, with little effect on secondary structure.[47]

V. REVERSIBILITY OF UNFOLDING TRANSITIONS: CLOSING REMARKS

The results in this work indicate that the protocol for ''centrifugal'' unfolding constitutes a robust and reproducible approach to denature lysozyme. Depending on the bath relaxation, we can either obtain fully unfolded conformers rapidly (when using a small τ) or an ensemble of partly unfolded conformers with longer persistence (when using a large τ). From a pragmatic point of view, the procedure is useful as a tool to generate reasonable unfolded conformations for a neutral protein. The computer-simulated ''denatured structures'' have common features with the available experimental data on the unfolded state.[12] As well, they provide convenient starting points for studies of relaxation and refolding.[29–32] It remains to be seen, however, whether protein unfolding by fast spinning is physically achievable in practice in the gas phase. In contrast, we conjecture that a similar process may be unlikely in solution given the need of an induction period of 500 ps of systematic (nonrandom) rotations for unfolding to occur.

The same conditions for unfolding apply even if one starts from other initial conformers. As an illustration, let us consider briefly the relaxation behavior of an initially unfolded structure, under conditions of strong bath coupling. Figure 9 shows two MD trajectories that start from the same fully unfolded conformer, selected from an initial run using $\tau=10$ fs. In this case, we have rerandomized the initial distribution of velocities, reimposed *in vacuo* boundary conditions with $\tau=10$ fs, and continued the run for 2 ns. The first 500 ps of the trajectories are represented in thick line, whereas the last 1500 ps are depicted in thin line. In both runs, the initial relaxation behavior leads to compactization. During the first 500 ps, the results resemble those observed in weak bath couplings (e.g., $\tau=100$ fs).[29–32] However, these compact intermediates eventually *reunfold* when using a $\tau=10$ fs coupling. As shown in Fig. 9, reunfolding leads to structures similar to the initial transient, even though the ''return'' path is different. These results suggest that our technique can cause unfolding in compact lysozyme structures, *whether native or not*. Cycles of unfolding and refolding are

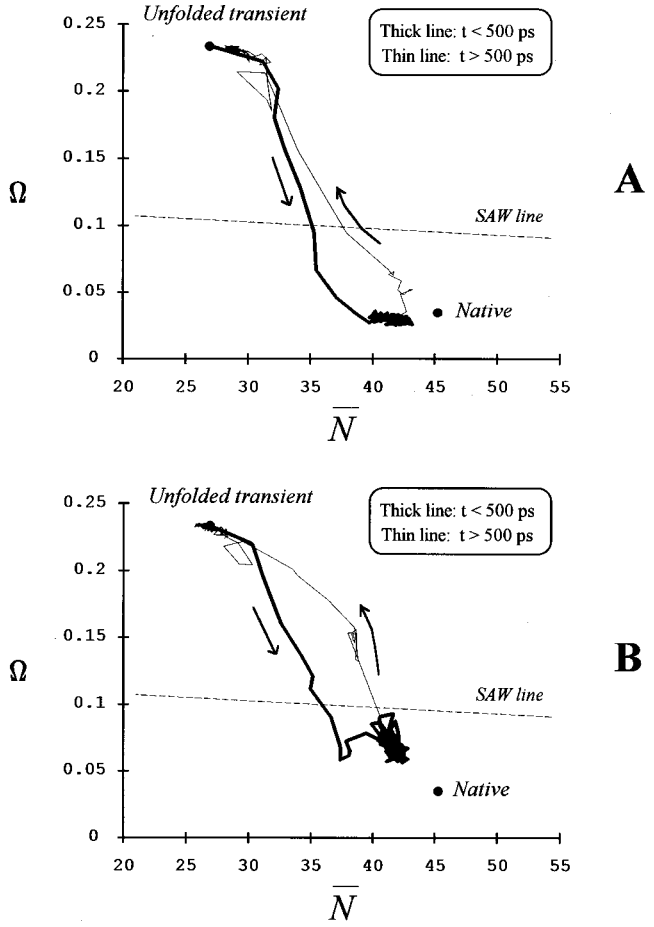

FIG. 9. Two examples of a reunfolding transition following MD relaxation from an unfolded transient. In these simulations, the initial structure is unfolded and subject to relaxation while using a strong coupling to the thermal bath ($\tau = 10$ fs). The results for the 2-ns MD run are separated into the first 500 ps (thick lines) and the last 1500 ps (thin lines). In both cases, the initial behavior is dominated by chain collapse as Ω decreases faster than the increase in $\overline{N}$. Eventually, reunfolding occurs by following a different pathway. The results show that cycles of unfolding–refolding are possible, but not strictly by reversible paths.

also possible, although not by reversed paths. Such differences in unfolding and refolding pathways are also believed to be the rule for processes in solution.[34,35,47]

The results in Fig. 9 also suggest that the centrifugal unfolding process provides a robust representation for the corresponding unfolded continuum. As indicated by all shape descriptors, one obtains the same ensemble of denatured conformers by direct unfolding or by relaxation followed by re-unfolding. In other words, both compact and noncompact conformers are *dynamically connected* in a reversible manner when using this procedure. From the practical viewpoint, this result is important because it confirms that the elongated lysozyme conformers are reproducible structures, typical of the unfolded continuum. Therefore, they represent reasonable choices for studies of relaxation dynamics in denatured lysozyme.[29–32] Morever, the results are also conceptually important, because they provide a stringent criterion to assess how reliably the denatured state is represented. Often, MD simulations of relaxation start from an *arbitrary* nonnative structure, believed to be typical of the denatured state. However, not all denaturing conditions are equivalent and structures from different unfolding continua may not even be dynamically connected.[34] Consequently, an incorrect choice for the initial conformer can lead to relaxation paths that do not represent a meaningful refolding process under the simulation conditions. We believe that the choice of initial unfolded transient should be checked and justified. Our proposal is to use the fact that the same unfolded conformers must occur in cycles of folding–unfolding (Fig. 9) as an operational definition for a ''stable'' computer-simulated denatured state.

In closing, it should be said again that we view the ''centrifugal'' effects associated with fast spinning only as a tool to bring the molecule to a state where it responds to the action of weak forces by unfolding. As a physical phenomenon, one can visualize this process as a possible behavior *in vacuo*. Yet, the important result is that the approach serves as a computational tool to uncover a structural response that may be intrinsic to the lysozyme fold.

It must be noted that we have no reason to believe that fast spinning should always elicit unfolding. The ability to activate rotationally an unfolding mode is expected to depend on protein tertiary structure. Current work in our laboratory on other proteins will help us clarify which tertiary structural features are key to facilitate (or favor) centrifugal unfolding.

ACKNOWLEDGMENTS

We thank Professor Curt Reimann (Lund) for valuable discussions and Naomi Grant (Sudbury) for her comments on the manuscript. G.A.A. acknowledges support from NSERC (Canada) and O.T. is grateful to NFR (Swedish Natural Sciences Council) for financial support.

[1] K. Shelimov and M. F. Jarrold, J. Am. Chem. Soc. **118**, 10313 (1996).
[2] D. S. Gross, P. D. Schnier, S. E. Rodríguez-Cruz, C. K. Fagerquist, and E. K. Williams, Proc. Natl. Acad. Sci. U.S.A. **93**, 3143 (1996).
[3] K. B. Shelimov, D. E. Clemmer, R. R. Hudgins, and M. F. Jarrold, J. Am. Chem. Soc. **119**, 2240 (1997).
[4] J. S. Valentine, J. G. Anderson, A. D. Ellington, and D. E. Clemmer, J. Phys. Chem. B **101**, 3891 (1997).
[5] F. W. McLafferty, Z. Guan, U. Haupts, T. D. Wood, and N. L. Kelleher, J. Am. Chem. Soc. **120**, 4732 (1998).
[6] S. W. Englander, Annu. Rev. Biophys. Biomol. Struct. **29**, 213 (2000).
[7] N. L. Kelleher, H. Y. Lin, G. A. Valaskovic, D. J. Aaserud, E. K. Fridriksson, and F. W. McLafferty, J. Am. Chem. Soc. **121**, 806 (1999).
[8] M. F. Jarrold, Acc. Chem. Res. **32**, 360 (1999).
[9] C. S. Hoaglund-Hyzer, A. E. Counterman, and D. E. Clemmer, Chem. Rev. **99**, 3027 (1999).
[10] M. F. Jarrold, Annu. Rev. Phys. Chem. **51**, 170 (2000).
[11] P. A. Sullivan, C. T. Reimann, J. Axelsson, S. Altmann, A. P. Quist, and B. U. R. Sundqvist, J. Am. Soc. Mass Spectrom. **7**, 329 (1996).
[12] C. T. Reimann, P. A. Sullivan, J. Axelsson, A. P. Quist, S. Altman, P. Roepstorff, I. Velázquez, and O. Tapia, J. Am. Chem. Soc. **120**, 7608 (1998).
[13] P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **92**, 2426 (1995).
[14] C. T. Reimann, I. Velázquez, and O. Tapia, J. Phys. Chem. B **102**, 9344 (1998).
[15] P. L. Privalov, in: T. E. Creighton (ed.), Protein Folding (W.H. Freeman, New York, 1992).
[16] D. Alonso and V. Daggett, Protein Sci. **7**, 860 (1998).
[17] S. L. Kazmirski and V. Daggett, J. Mol. Biol. **277**, 487 (1998).
[18] P. Bruscolini and L. Casetti, Phys. Rev. E **61**, R2208 (2000).
[19] M. Marchi and P. Ballone, J. Chem. Phys. **110**, 3697 (1999).
[20] P. Ferrara, J. Apostolakis, and A. Caflisch, J. Phys. Chem. B **104**, 4511 (2000).

[21] M. Rief, F. Oesterhelt, B. Heymann, and H. E. Graub, Science **275**, 1297 (1997).
[22] M. S. Z. Kellermayer, S. B. Smith, H. L. Granzier, and C. Bustamante, Science **276**, 1112 (1997).
[23] P. E. Marszalek, H. Lu, H. Li, M. Carrión-Vázquez, A. F. Oberhauser, K. Schulten, and J. M. Fernández, Nature (London) **402**, 100 (1999).
[24] M. N. Tamashiro and P. Pincus, Phys. Rev. E **63**, 021909 (2001).
[25] B. Isralewitz, M. Gao, and K. Schulten, Curr. Opin. Struct. Biol. **11**, 224 (2001).
[26] M. Shtilerman, G. H. Lorimer, and S. W. Englander, Science **284**, 822 (1999).
[27] V. Grantcharova, E. J. Alm, D. Baker, and A. L. Horwich, Curr. Opin. Struct. Biol. **11**, 70 (2001).
[28] C. T. Reimann, I. Velázquez, and O. Tapia, J. Phys. Chem. B **102**, 2277 (1998).
[29] I. Velázquez, C. T. Reimann, and O. Tapia, J. Am. Chem. Soc. **121**, 11468 (1999).
[30] G. A. Arteca, I. Velázquez, C. T. Reimann, and O. Tapia, Phys. Rev. E **59**, 5981 (1999).
[31] G. A. Arteca, I. Velázquez, C. T. Reimann, and O. Tapia, J. Chem. Phys. **111**, 4774 (1999).
[32] G. A. Arteca, M. Paulino, C. T. Reimann, and O. Tapia, Phys. Chem. Chem. Phys. **2**, 5259 (2000).
[33] G. A. Arteca, C. T. Reimann, and O. Tapia, J. Phys. Chem. B **104**, 11360 (2000).
[34] A. V. Finkelstein, Protein Eng. **10**, 843 (1997).
[35] A. R. Dinner and M. Karplus, J. Mol. Biol. **292**, 403 (1999).
[36] A. Miranker, C. V. Robinson, S. E. Radford, R. T. Aplin, and C. M. Dobson, Science **262**, 896 (1993).
[37] W. F. van Gunsteren and H. J. C. Berendsen, *Groningen Molecular Simulation (GROMOS) Library Manual* (Biomos, Groningen, 1987).
[38] J. Åqvist, W. F. van Gunsteren, M. Leijonmarck, and O. Tapia, J. Mol. Biol. **183**, 461 (1985).
[39] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984).
[40] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1989).
[41] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic, San Diego, 1996).
[42] G. A. Arteca and O. Tapia, J. Mol. Graphics Modell. **19**, 102 (2001).
[43] R. G. Littlejohn and M. Reinsch, Rev. Mod. Phys. **69**, 213 (1997).
[44] J. Karczmarek, J. Wright, P. Corkum, and M. Ivanov, Phys. Rev. Lett. **82**, 3420 (1999).
[45] D. M. Villeneuve, S. A. Aseyev, P. Dietrich, M. Spanner, M. Yu. Ivanov, and P. B. Corkum, Phys. Rev. Lett. **85**, 542 (2000).
[46] M. Gupta and R. N. Zare, Nature (London) **407**, 33 (2000).
[47] U. Mayor, C. M. Johnson, V. Daggett, and A. R. Fersht, Proc. Natl. Acad. Sci. U.S.A. **97**, 13518 (2000).
[48] G. A. Arteca and P. G. Mezey, Biopolymers **32**, 1609 (1992).
[49] E. J. Janse van Rensburg, D. W. Sumners, E. Wasserman, and S. G. Whittington, J. Phys. A **25**, 6557 (1992).
[50] G. A. Arteca, Biopolymers **33**, 1829 (1993).
[51] G. A. Arteca, O. Nilsson, and O. Tapia, J. Mol. Graph. **11**, 193 (1993).
[52] V. Katritch, J. Bednar, D. Michoud, R. G. Scharein, J. Dubochet, and A. Stasiak, Nature (London) **384**, 142 (1996).
[53] A. V. Vologodskii, N. J. Crisona, B. Laurie, P. Pieranski, V. Katritch, J. Dubochet, and A. Stasiak, J. Mol. Biol. **278**, 1 (1998).
[54] G. A. Arteca, J. Chem. Inf. Comput. Sci. **39**, 550 (1999).
[55] G. A. Arteca and O. Tapia, Int. J. Quantum Chem. **80**, 848 (2000).
[56] A. Baumgärtner, J. Chem. Phys. **98**, 7496 (1993).
[57] R. Koradi, M. Billeter, and K. Wüthrich, J. Mol. Graph. **14**, 51 (1996).
[58] G. A. Arteca, I. Velázquez, C. T. Reimann, and O. Tapia, Chem. Phys. Lett. **327**, 245 (2000).
[59] D. K. Klimov and D. Thirumalai, J. Chem. Phys. **109**, 4119 (1998).
[60] D. Thirumalai and D. K. Klimov, Curr. Opin. Struct. Biol. **9**, 197 (1999).