

# Pathways for folding and re-unfolding transitions in denatured conformations of anhydrous proteins

Gustavo A. Arteca<sup>a,\*</sup>, Kalyani Veluri<sup>a</sup>, O. Tapia<sup>b</sup>

<sup>a</sup> *Département de Chimie et Biochimie, Laurentian University, Sudbury, Ont., Canada P3E 2C6*

<sup>b</sup> *Department of Physical Chemistry, Uppsala University, Box 579, S-751 23 Uppsala, Sweden*

Received 23 August 2002

## Abstract

Using molecular dynamics simulations, we explore the mechanism that may underlie the type of folding transitions observed experimentally in anhydrous proteins. Recent relaxation studies of a partly unfolded  $\alpha$ -helical bundle suggest an approach to simulating cycles of folding/re-unfolding transitions. Here, we provide evidence that such processes may be general, by uncovering a similar relaxation pattern in lysozyme, a more flexible  $\alpha/\beta$  protein. We find that, while refolding is initiated by partial polymer collapse, the reversible re-unfolding of compact nonnative structures proceeds as a correlated transition not unlike the direct unfolding of lysozyme from its native state.

© 2003 Elsevier Science B.V. All rights reserved.

## 1. Introduction

Experiments in mass spectrometry show that anhydrous proteins diffusing in a low-pressure inert gas exhibit folding–unfolding transitions that resemble, to some extent, those observed in solution [1–3]. Even though we do not know yet the extent to which these phenomena are pertinent to cellular proteins, it is becoming clear that ‘folding’ is a process intrinsic to the protein backbone, one with no obligatory mediation of water. Details of these transformations and the factors that control them are still beyond experimental reach; however, computer simulations begin to provide glimpses of

the basic principles at play [4,5]. Here, we focus on the pattern of paths associated with *simulated cycles of folding–unfolding transitions in anhydrous proteins* with unhindered translations and rotations (i.e., proteins in vacuo).

Under simulation conditions that allow inertial spinning, it is possible to unfold a globally neutral protein even at room temperature [6,7]. Within this process, the unfolding bias can be regulated using the *bath coupling*, i.e., through energy exchange between the molecules vibro-rotations and a thermal bath. The resulting process has common features with the unfolding of dehydrated protein ions within a drift tube, including the formation of elongated conformations with residual secondary structure [1–3,8]. Regarding the unfolding mechanism, computer simulations suggest pathways where the loss of chain compactness correlates

\* Corresponding author. Fax: +1-705-675-4844.

E-mail address: [gustavo@laurentienne.ca](mailto:gustavo@laurentienne.ca) (G.A. Arteca).

with a simplification in folding complexity [7], a behaviour also observed in the unfolding of protein ions with near-critical charge [9]. In contrast, highly charged ions appear to unfold initially with constant compactness [9,10]. By assuming that these structures are representative of the denatured state, recent work has begun to explore the relaxation pathways leading to compact, possibly native-like, folded conformers [4,5,11]. In particular, we have recently shown that one can simulate cyclic processes whereby a partly unfolded conformer of a stiff  $\alpha$ -helical bundle may first fold, and then re-unfold to the *same* denatured state [12]. Here, we extend this previous work, and apply it to study the detailed relaxation behaviour of fully unfolded lysozyme, which is a more flexible  $\alpha/\beta$  protein. We first summarize the simulation and analysis methods, which rely on trajectory ensembles. Particular attention is given to following large-scale molecular shape features that distinguish between folding and re-unfolding paths. We discuss differences among paths according to polymer collapse and chain entanglement. Finally, we show how the relative dominance of re-unfolding over folding transitions can be modified smoothly by changing the coupling to the simulated thermal bath.

## 2. Simulation protocol and fold characterization

We generate ensembles of constant-temperature molecular dynamics (MD) trajectories for a globally neutral protein under in vacuo boundary conditions. Charges are distributed over the molecule, but each amino acid residue is *locally neutral*. The approach is meant to mimic the behaviour of a protein at neutral pH, and not zwitterions. In addition, we introduce boundary conditions whereby inertial rotation can persist over time, as it would in actual dehydrated proteins in flight. As a result of the onset of global spinning, energy transfer between rotations and extended collective vibrations result in ‘*centrifugal unfolding*’ [6]. The degree of unfolding and the time scale can be regulated by changing the constant  $\tau$  that couples the molecule with a simulated thermal bath [7,13]. Here, we use the Berendsen thermostat

model [14], where the instantaneous ‘molecular temperature’  $T$  drifts toward the bath temperature  $T_0$  according to a cooling law:  $dT/dt = (T_0 - T)/\tau$ . If the coupling is strong (i.e., small  $\tau$ ), unfolding is fast and predominant. A weaker coupling can still elicit unfolding by similar pathways, although at a lower rate. The bath model provides a soft mechanism to induce configurational changes by non-equilibrium dynamics. Using in vacuo simulations, it is possible to monitor the early steps of folding–unfolding transitions with nanosecond-long trajectories. Here, we adapt this simulation protocol to study systematically the effect of the strength of the bath coupling over the relaxation behaviour of an unfolded lysozyme conformer.

Simulations are carried on hen egg-white lysozyme (PDB code *1hel*), modelled with the neutral-protein parameter set 37D4 of the Gromos force field [15,16]. Four disulfide bridges are maintained during all MD runs at  $T_0 = 293$  K. Relaxation studies start from a fully unfolded conformer of lysozyme. This initial structure is selected from a generic unfolding run, obtained after coupling strongly the native state of lysozyme to the simulated thermal bath ( $\tau = 0.01$  ps). This unfolding run is representative of the behaviour for all strong-coupling simulations in [7]. The initial conformer has a residual secondary structure, and its size is comparable to the experimental values observed in unfolded dehydrated lysozyme. (See a snapshot in Fig. 1.)

Independent relaxation trajectories are generated by changing the initial distribution of random velocities for  $T_0 = 293$  K. We have collected data from twenty 2 ns-long runs for each of five  $\tau$  values: 0.01, 0.04, 0.08, 0.12, and 0.16 ps, with integration step  $\Delta t = 1$  fs. Previously, we discussed the effects of bath coupling on lysozyme *unfolding* [7]. Now, we focus on how this affects *relaxation and refolding*. The same set of 20 random number seeds are used for each trajectory series; all effects observed are thus due only to a change in relaxation constant.

Taking snapshots every 5 ps, we have monitored the evolution of various molecular shape properties. As proposed elsewhere [4,5], we focus on descriptors that convey the extent of polymer collapse and a degree of chain folding. The inter-

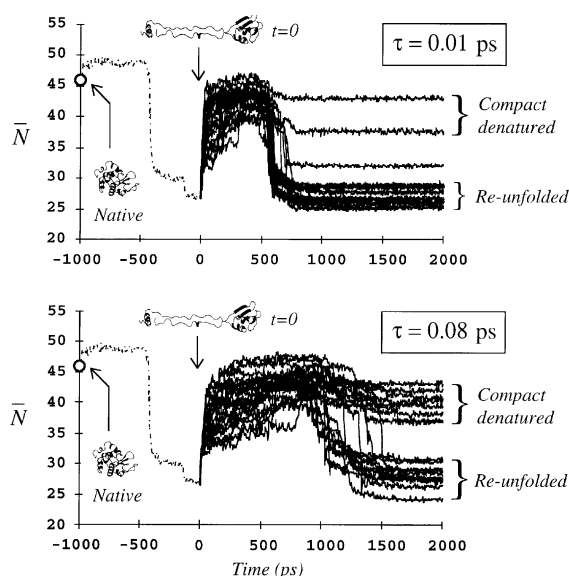


Fig. 1. Evolution of chain entanglement during the relaxation behaviour of unfolded *1hel* lysozyme. The initial unfolding run is in dashed line; the snapshot at  $t = 0$  is the starting structure for relaxation. When using a strong coupling, re-unfolding is both rapid and prevalent ( $\tau = 0.01$  ps, top diagram). When the coupling is weaker, re-unfolding is delayed and competes with the formation of compact denatured intermediates ( $\tau = 0.08$  ps, bottom diagram).

play of these competing properties is central for understanding protein folding and unfolding mechanisms [17,18].

The compactness of the  $\alpha$ -carbon backbone can be conveyed simply by the *asphericity*,  $\Omega$ , defined as:  $\Omega = \{(I_1 - I_2)^2 + (I_1 - I_3)^2 + (I_3 - I_2)^2\} / 2(I_1 + I_2 + I_3)^2$ , with the principal moments of inertia  $\{I_i\}$ . For a 3D-chain, the asphericity varies between  $\Omega \approx 0$  in a compact globule, to  $\Omega \approx 1/4$  for elongated rod-like chains. The configurationally averaged value in self-avoiding walks with excluded-volume interaction,  $\langle \Omega \rangle_{\text{SAW}}$ , provides a measure of compactness that separates collapsed conformers (where  $\Omega < \langle \Omega \rangle_{\text{SAW}}$ ) from elongated ones (where  $\Omega > \langle \Omega \rangle_{\text{SAW}}$ ) [4].

Folding complexity relays the spatial organization of the backbone trace. To this end, we use descriptors of *chain entanglement* derived from the overcrossing probability distribution  $\{A_N^{(K)}\}$ , which depends on both the protein geometry and the backbone connectivity [19]. The distribution

$\{A_N^{(K)}\}$  measures the probability of observing  $N$  projected bond–bond crossings (or ‘overcrossings’), for an arbitrary projection of the backbone in a rigid configuration  $K$ . From  $\{A_N^{(K)}\}$ , two properties can provide a summary description of *global folding features* that is independent of the definition of secondary and tertiary structure. These are the first moment of the distribution (the *mean overcrossing number*,  $\bar{N}$ ) and the root-mean-square deviation in overcrossing probability,  $\sigma = \{\sum_{N \geq 0} (A_N^{(K)} - A_N^{(K*)})^2\}^{1/2}$ , defined with respect to a reference structure  $K^*$ . In our case,  $K^*$  is the crystal structure of *1hel* lysozyme. During an unfolding transition, we expect  $\bar{N}$  to decrease up to a value  $\bar{N} = 0$  in rod-like chains. As well, when comparing compact and noncompact structures, we expect  $\sigma \in [0, 1]$  [4]. Both descriptors are translationally and rotational invariant;  $\sigma$  does not require molecular fitting. Regarding the numerical (and possibly analytical) computation of  $\{A_N^{(K)}\}$  and  $\bar{N}$ , see [19,20].

For our present goals, the  $(\bar{N}, \Omega)$  shape space provides a simple representation to understand how polymer collapse and folding features relate to each other during molecular dynamics. In the next sections, we discuss the evolution of these descriptors and the pattern of transitions in  $(\bar{N}, \Omega)$ -space. As well, we discuss how the choice of relaxation constant  $\tau$  affects the population of folded, unfolded, and intermediate conformers along MD trajectories.

### 3. Effect of bath coupling on conformer population during relaxation

In the crystal, the native state of lysozyme has entanglement  $\bar{N} \approx 45.5$  and asphericity  $\Omega \approx 0.035$  [21]. Once exposed to the unfolding bias of a strong bath coupling ( $\tau = 0.01$  ps), lysozyme normally undergoes centrifugal denaturation within 700 ps [7], thereby reducing the  $\bar{N}$  values. A typical 1 ns-long MD run for unfolding is shown in Fig. 1 (dashed line). The structure at the end of this trajectory becomes the initial conformer for relaxation (snapshot at  $t = 0$  in Fig. 1). Within the precision of the numerically computed entanglement descriptor, we find  $\bar{N} = 26 \pm 1$  for this

structure. (All snapshots were created with the program Molmol [22].)

Fig. 1 illustrates the key differences in relaxation behaviour depending on the bath coupling. The top diagram shows the mean response emerging from 20 MD runs at strong coupling,  $\tau = 0.01$  ps; the bottom diagram deals with a weaker coupling,  $\tau = 0.08$  ps. In both cases, the immediate response of the unfolded structure is to increase the mean overcrossing number through the onset of *partial* compaction [21]. However, clear differences do occur later:

- Under strong coupling, entangled intermediates appear during the first 500 ps of relaxation, corresponding to the formation of compact (denatured) conformers. Most of these structures *re-unfold* at  $t \approx 600$  ps, leading to  $\bar{N}$  values similar to those for the conformer at  $t = 0$ . Only a small subset of trajectories remain at the entanglement level for compact denatured conformers.
- Under a weaker coupling, compact and semi-compact intermediates persist for longer times (up to  $t \approx 1000$  ps after relaxation). However, we find that re-unfolding is still possible under these conditions, although this occurs in a smaller fraction of the trajectories.

Table 1 completes the trend illustrated in Fig. 1. Here, we list the fraction of trajectories that proceed to re-unfold. To qualify for ‘re-unfolding,’ we tested two criteria. In one case, we use  $\bar{N} < 30$ , a

Table 1  
Fraction of trajectories that proceed to re-unfold

$\tau$ (ps)	% Re-unfolded ( $\bar{N} < 30$ )	% Re-unfolded ( $\Omega > 0.15$ )
0.01	85	90
0.04	65	90
0.08	50	60
0.12	65	75
0.16	60	75

For each coupling constant  $\tau$ , results are derived from the conformers at the end of 20 trajectories ( $t = 2$  ns). Two criteria for ‘full unfolding’ are compared, based on measuring chain entanglements ( $\bar{N}$ ) or anisometry ( $\Omega$ ). The main difference appears at  $\tau = 0.04$  ps, which produces swollen conformers ( $\Omega > 0.15$ ) with still significant entanglement ( $\bar{N} > 30$ ). Many unfolded conformers resemble the initial denatured structure of lysozyme. See Table 2 for a quantitative assessment of similarity.

range consistent with the initial conformer for relaxation. On the last column, we define the re-unfolded states as those with a level of compactness below that for swollen random walks with the same contour length,  $\Omega > 0.15$  [21]. Both definitions yield a similar behaviour: re-unfolding is always possible, but becomes a less dominant outcome as the coupling weakens. This result is consistent with the fact that the angular velocity associated with global spinning decreases as  $\tau$  grows, thus diminishing the energy transfer between rotations and unfolding vibrations. (Regarding the deviation for  $\tau = 0.04$  ps, see below.)

Fig. 2 gives the population histograms for conformers obtained at the end of each trajectory. For a better comparison, we consider two groups of 40 structures at  $t = 2000$  ps: one for strong-coupling ( $\tau = 0.01$  and 0.04 ps) and another for weak ones ( $\tau = 0.08$  and 0.12 ps). It is clear that the main difference among relaxation patterns is in the *proportion* of folded vs unfolded conformations; *the folding features observed appear to be the same for all  $\tau$  values*. This is consistent with observations for cytochrome *c'* [12], even though the configurational transitions in the latter are very different from those of lysozyme.

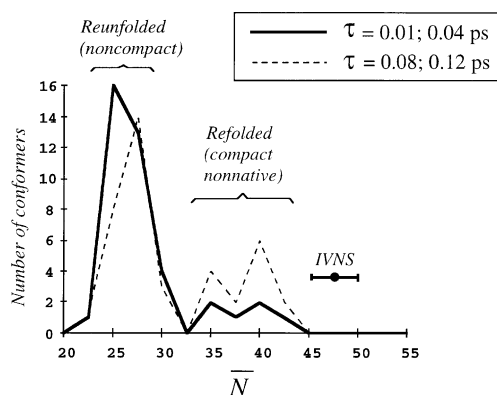


Fig. 2. Population histogram, in terms of chain entanglement, corresponding to the distribution of conformers at  $t = 2$  ns. The thick line collects the data for stronger couplings and the dashed line for weaker ones. Each histogram contains forty conformers. The error bar is the mean overcrossing number for the metastable in vacuo native structure. Note that  $\tau$  affects the relative conformer population but not the shape features. The ensemble of noncompact re-unfolded structures contains many that are indistinguishable from the one at  $t = 0$ .

We can check our inferences from Fig. 2 by monitoring also the *relative* overcrossing descriptor  $\sigma$ . This descriptor compares the folding features of instantaneous configurations with those of lysozyme in the crystal. The result is a more strict criterion for shape comparison, given that conformers with similar  $\bar{N}$  value can still differ in  $\sigma$  value.

At  $t = 0$ , the initial structure for relaxation is quite different from the native state,  $\sigma = 0.22 \pm 0.01$ . (Note that relative differences among mutations of the native state of lysozyme amount to only  $\sigma = 0.03 \pm 0.02$ .) Table 2 lists the results for the mean  $\sigma$  values computed from the 20 structures observed at  $t = 2000$  ps for each  $\tau$ . In order to convey effectively any *shape* differences among  $\sigma$ -populations, we also list the kurtosis ( $\kappa$ ) of the conformer histograms. The kurtosis, which is the *fourth* moment of the distribution, describes its *flatness*. When  $\kappa > 0$ , we expect a sharp single-peak function; a negative  $\kappa$  value is consistent with a flatter distribution, possibly with several comparable smaller maxima. From Table 2, we observe:

(a) The relative fold similarity returns to values close to the initial one ( $\sigma \approx 0.22$ ) for  $t > 1000$  ps in  $\tau = 0.01$  ps and  $t > 1500$  ps in  $\tau = 0.04$  ps. These results are consistent with those for  $\bar{N}$  in Table 1.

- (b) With weaker couplings ( $\tau > 0.04$  ps),  $\sigma$  increases along the MD run, but remains smaller than the value at  $t = 0$ . This is consistent with having a fraction of conformers that are not unfolded.
- (c) The  $\sigma$ -distribution kurtosis increases in trajectories with the stronger *and* the weaker couplings. This increase in  $\kappa$  suggests *single* populations of conformers. At  $\tau = 0.01$  ps, the dominant conformers should all be re-unfolded; at  $\tau = 0.16$  ps, they should be semi-compact.
- (d) In contrast with (c),  $\kappa$  remains negative for most of the trajectories at intermediate  $\tau$  values. In particular, the  $\sigma$ -distributions are quite flat for  $\tau = 0.08$  ps and  $\tau = 0.12$  ps. This result indicates that a range of conformers coexist, spanning  $\sigma$  values between those consistent with compact denatured structures ( $\sigma < 0.1$ ) up to those characteristic of re-unfolded conformers ( $\sigma > 0.2$ ).

In summary, Table 2 results are consistent with the notion that a change in  $\tau$  shifts the conformer population, but *does not* produce new folding features. The fact that both  $\bar{N}$  and  $\sigma$  return to their initial values confirms that a reversible re-unfolding transition occurred.

Table 2

Distribution of root-mean-square deviations in overcrossing probabilities ( $\sigma$ ) for relaxation constants ( $\tau$ ) and times ( $t$ ) along the MD trajectories

$\tau$ (ps)	$t = 500$ (ps)	$t = 1000$ (ps)	$t = 1500$ (ps)	$t = 2000$ (ps)
0.01	0.112 $\pm$ 0.014 (0.065)	0.214 $\pm$ 0.013 (0.707)	0.215 $\pm$ 0.012 (1.218)	0.216 $\pm$ 0.012 (1.407)
0.04	0.119 $\pm$ 0.010 (-1.104)	0.176 $\pm$ 0.015 (-0.528)	0.203 $\pm$ 0.013 (-0.314)	0.205 $\pm$ 0.015 (-0.210)
0.08	0.118 $\pm$ 0.014 (0.810)	0.125 $\pm$ 0.015 (-1.069)	0.181 $\pm$ 0.020 (-1.608)	0.184 $\pm$ 0.018 (-1.675)
0.12	0.106 $\pm$ 0.012 (-0.484)	0.108 $\pm$ 0.012 (-1.107)	0.182 $\pm$ 0.019 (-1.219)	0.198 $\pm$ 0.021 (-0.898)
0.16	0.105 $\pm$ 0.018 (-0.562)	0.104 $\pm$ 0.020 (-0.154)	0.178 $\pm$ 0.026 (0.786)	0.189 $\pm$ 0.027 (1.776)

The first row for each  $\tau$  gives the mean  $\sigma$ , with 95% confidence, computed over 20 trajectories. Full re-unfolding to the original structure is marked in bold. Values in parenthesis below each  $\tau$  row indicate the kurtosis ( $\kappa$ ) of the  $\sigma$  distribution. Note that  $\kappa > 0$  values at the end of the trajectories for  $\tau = 0.01$  ps and  $\tau = 0.16$  ps suggest the formation of well-defined population of conformers (fully unfolded conformers and partly unfolded intermediates, respectively). In contrast,  $\kappa < 0$  values for intermediate  $\tau$  couplings are indicative of the co-existence of at least two significant conformer populations.

#### 4. Differences between folding and re-unfolding pathways

The results in Section 3 suggest that the present unfolding technique produces *dynamically connected* conformers. That is, folded and unfolded conformers occur in regions visited by reversible cycles of folding–unfolding transitions. Here we show that, in agreement with conjectures in the literature [23,24] and hints from limited computer experiments [4,7], folding and re-unfolding paths appear not to be equivalent.

In order to compare folding paths, we use the  $(\bar{N}, \Omega)$ -map for the evolution of chain entanglements and compactness. Fig. 3 is illustrative of the range of behaviours that are typically observed for all values of the coupling constant  $\tau$ , which regulates the energy exchange between the molecule and simulated thermal bath. The diagrams compare the first 500 ps of MD simulation (dashed line) with the last 1500 ps (thick line), in four trajectories with  $\tau = 0.04$  ps. The separated regions for compact and noncompact conformers are established unbiasedly using the mean shape properties of simple polymers with excluded volume interaction, as explained in [21].

The top diagram in Fig. 3 shows the more common behaviour, with distinct paths for compaction and re-unfolding. For example, the path section (2), corresponding to  $0 \leq t < 500$  ps, is initiated by a rapid reduction in  $\Omega$  (possibly, polymer collapse), followed by an oscillation in  $\Omega$  that suggests re-expansion of a compact chain. The continuation section (2'), in contrast, shows a softer increase in  $\Omega$  at the expense of  $\bar{N}$  something that is indicative of a re-unfolding transition where compactness and tertiary fold are more correlated.

The bottom diagram in Fig. 3 shows that other, distinct behaviours are also possible, even if they are less likely. The trajectory sections denoted by (3) and (3') serve to demonstrate that *near-reversibility* between the initial folding path and late re-unfolding path can occasionally be observed. Finally, the trajectory sections (4) and (4') show that compaction may also yield long-lived intermediates where re-unfolding is delayed or altogether frustrated. These types of behaviour appear to be less likely when using smaller  $\tau$  values.

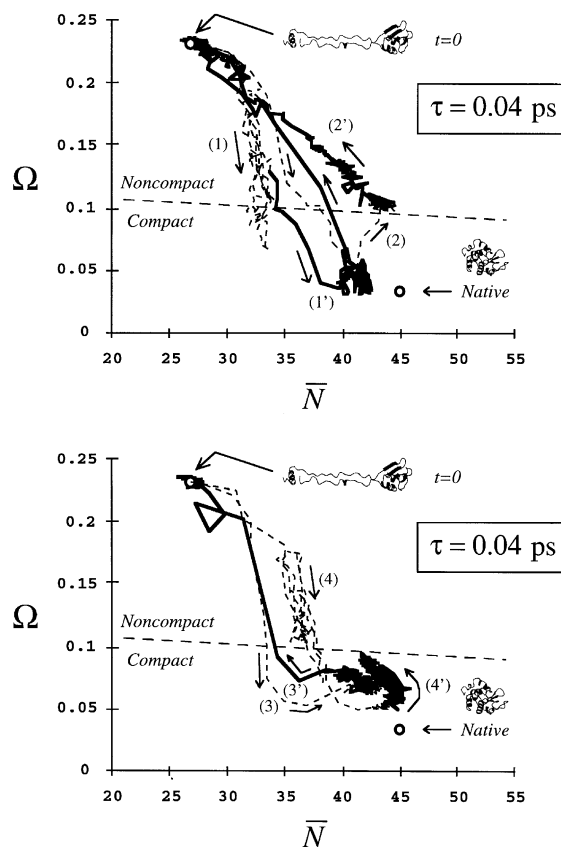


Fig. 3. Distinct relaxation trajectories in shape space for  $\tau = 0.04$  ps. The dashed lines denoted with plain numbers, (1)–(4), join conformers with  $0 \leq t < 500$  ps. The thick lines denoted with primes, (1')–(4'), correspond to  $500 \leq t \leq 2000$  ps. The top diagram shows the most common behaviour, where folding and re-unfolding follow distinct paths. The bottom diagram shows a less common alternative, where there is near reversibility between 'early' and 'late' sections of the trajectory (e.g., paths (3) and (3')). In the case of the sections (4) and (4'), re-unfolding along a new path appears to be frustrated.

The average difference between folding and re-unfolding pathways is well illustrated by Fig. 4, corresponding to the MD trajectories with  $\tau = 0.16$  ps. Here, we define a 'folding path' as the section of a trajectory in  $(\bar{N}, \Omega)$ -space that, starting from the initial structure, can reach a long-lived compact transient *before* re-unfolding sets in. In other words, any trajectory not producing a compact intermediate is not counted as a 'folding path,' regardless of whether it remains frustrated or unfold later. (This is the same notion used to

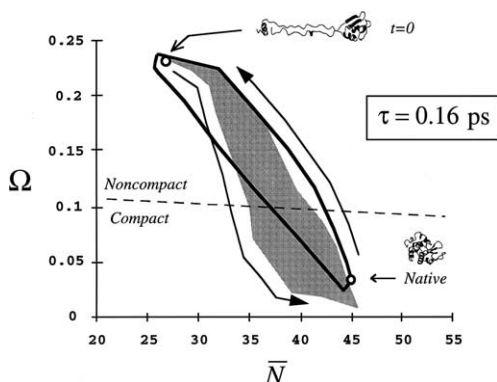


Fig. 4. Regions in  $(\bar{N}, \Omega)$ -space associated with folding and re-unfolding paths for  $\tau = 0.16$  ps. The grey area spans the average of *folding paths*, i.e. the trajectory sections leading from the initial structure to a compact intermediate. (Trajectories not leading to compaction are excluded.) The region enclosed in thick line contains the average of *re-unfolding paths*, i.e., the trajectory sections leading from a compact intermediate to a noncompact structure. The arrows indicate the time evolution. Note that folding and unfolding do not follow equivalent paths.

define the *folding probability* of a chain in [17]. For  $\tau = 0.16$  ps, we find a 70% folding probability.) Typically, the folding paths include the first 50–150 ps of a trajectory. The compact intermediates that follow them may last up to 1000 ps. The average region spanned by these folding paths defined appear as a grey area in Fig. 4. Similarly, we define *re-unfolding paths* as the trajectory sections in  $(\bar{N}, \Omega)$ -space leading to a noncompact structure from a compact intermediate. The average region spanned by these trajectories is enclosed inside the thick line in Fig. 4. As the figure shows, there is a clear difference between the two families of paths. The folding region shows a faster decrease in  $\Omega$  compared to the increase in  $\bar{N}$  (note the time arrows in Fig. 4). In contrast, the re-unfolding region correlates these two properties more evenly. On average, there is no microscopic reversibility between folding and re-unfolding paths.

Note that the region corresponding to re-unfolding paths includes the neighbourhood of the  $t = 0$  structure. This finding confirming that reversible re-unfolding takes place even when using a weak bath coupling. On the other hand, the folding region marginally approaches the neighbourhood of the native state. Evidently, it is not possible to achieve native-like folding within 2 ns,

when starting from a fully unfolded lysozyme conformer under conditions that allow eventual re-unfolding. Before folding proceeds any further, a re-unfolding transition to the denatured state can take place in vacuo.

## 5. Closing remarks

In this work, we have shown that it is possible to connect reversibly structures belonging to the configurational continuum of a denatured protein. By varying the value of the bath coupling, i.e., through different rates of energy transfer within the molecule, it is possible to generate fast cycles of folding–unfolding transitions or long-lived compact intermediates. In either case, the simulation protocol is able to eventually return the protein to the initial structure used for relaxation. We believe that this dynamic connectivity and reversibility is an essential feature that one should always verify in MD simulations in order to ensure that the set of unfolded conformers is physically meaningful. We now have found that a similar reversibility occurs in a flexible protein such as lysozyme and in a stiffer one such as cytochrome *c'* [12]. This may suggest that the technique of centrifugal unfolding by bath-coupling regulation in vacuo is a general and robust tool to explore the qualitative pattern of configurational transitions during protein relaxation.

The findings in Section 4 fit well with the emerging picture, whereby the early steps of *folding* are dominated by a *partial compaction* of the chain, followed by a slower development of chain entanglement up to a level similar to that of the native state [4,5,21,25]. This behaviour is facilitated by oscillations in compactness ('chain breathing') that are only possible in a chain that is neither maximally compact nor configurationally frustrated. Under the present in vacuo boundary conditions and bias toward denaturing, re-unfolding can take place from a compact intermediate. When contrasting the results for re-unfolding in Fig. 4 with those in [7] for the unfolding of *native-like* lysozyme, we observe similar correlation between  $\bar{N}$  and  $\Omega$ , although shifted toward smaller  $\bar{N}$  values in Fig. 4. This result suggests that unfolding pathways may share some essential

characteristics that persist *whether the structure to be unfolded is native-like or a compact nonnative intermediate*.

## Acknowledgements

G.A.A. thanks the Department of Physical Chemistry (Uppsala University) for its hospitality. This research was supported by NSERC (Canada) and the Canada Research Chairs' Program.

## References

- [1] C.S. Hoaglund-Hyzer, A.E. Counterman, D.E. Clemmer, *Chem. Rev.* 99 (1999) 3037.
- [2] M.F. Jarrold, *Annu. Rev. Phys. Chem.* 51 (2000) 179.
- [3] E.R. Badman, C.S. Hoaglund-Hyzer, D.E. Clemmer, *Anal. Chem.* 73 (2001) 6000.
- [4] G.A. Arteca, O. Tapia, *J. Mol. Graph. Mod.* 19 (2001) 102.
- [5] G.A. Arteca, C.T. Reimann, O. Tapia, *Mass Spectrom. Rev.* 20 (2001) 402.
- [6] C.T. Reimann, I. Velázquez, O. Tapia, *J. Phys. Chem. B* 102 (1998) 2277.
- [7] G.A. Arteca, O. Tapia, *J. Chem. Phys.* 115 (2001) 10557.
- [8] C.T. Reimann, P.A. Sullivan, J. Axelsson, A.P. Quist, S. Altmann, P. Roepstorff, I. Velázquez, O. Tapia, *J. Am. Chem. Soc.* 120 (1998) 7608.
- [9] G.A. Arteca, C.T. Reimann, O. Tapia, *J. Phys. Chem. B* 105 (2001) 4992.
- [10] G.A. Arteca, I. Velázquez, C.T. Reimann, O. Tapia, *Chem. Phys. Lett.* 327 (2000) 245.
- [11] I. Velázquez, C.T. Reimann, O. Tapia, *J. Am. Chem. Soc.* 121 (1999) 11468.
- [12] G.A. Arteca, O. Tapia, *Chem. Phys. Lett.* 365 (2002) 148.
- [13] G.A. Arteca, O. Tapia, *J. Phys. Chem. B* 106 (2002) 1081.
- [14] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, *J. Chem. Phys.* 81 (1984) 3684.
- [15] J. Åqvist, W.F. van Gunsteren, M. Leijonmarck, O. Tapia, *J. Mol. Biol.* 183 (1985) 461.
- [16] W.F. van Gunsteren, H.J.C. Berendsen, *Groningen Molecular Simulation (Gromos) Library Manual*, Biomos, Groningen, 1987.
- [17] R. Du, V.S. Pande, A.Yu. Grosberg, T. Tanaka, E. Shakhnovich, *J. Chem. Phys.* 111 (1999) 10375.
- [18] J. Chahine, H. Nymeyer, V.B.P. Leite, N.D. Socchi, J.N. Onuchic, *Phys. Rev. Lett.* 88 (2002) 16101.
- [19] G.A. Arteca, *Bipolymers* 33 (1993) 1829.
- [20] G.A. Arteca, *J. Chem. Inf. Comput. Sci.* 39 (1999) 550.
- [21] G.A. Arteca, I. Velázquez, C.T. Reimann, O. Tapia, *Phys. Rev. E* 59 (1999) 5981.
- [22] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graph.* 14 (1996) 51.
- [23] A.V. Finkelstein, *Protein Eng.* 10 (1997) 843.
- [24] L. Mirny, E. Shakhnovich, *Annu. Rev. Biophys. Biomol. Struct.* 30 (2001) 361.
- [25] G.A. Arteca, I. Velázquez, C.T. Reimann, O. Tapia, *J. Chem. Phys.* 111 (1999) 4774.