

The Folding Transition-State Ensemble of a Four-Helix Bundle Protein: Helix Propensity as a Determinant and Macromolecular Crowding as a Probe

Hariato Tjong and Huan-Xiang Zhou*

Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida

ABSTRACT The four-helix bundle protein Rd-apocyt b_{562} , a redesigned stable variant of apocytochrome b_{562} , exhibits two-state folding kinetics. Its transition-state ensemble has been characterized by Φ -value analysis. To elucidate the molecular basis of the transition-state ensemble, we have carried out high-temperature molecular dynamics simulations of the unfolding process. In six parallel simulations, unfolding started with the melting of helix I and the C-terminal half of helix IV, and followed by helix III, the N-terminal half of helix IV and helix II. This ordered melting of the helices is consistent with the conclusion from native-state hydrogen exchange, and can be rationalized by differences in intrinsic helix propensity. Guided by experimental Φ -values, a putative transition-state ensemble was extracted from the simulations. The residue helical probabilities of this transition-state ensemble show good correlation with the Φ -values. To further validate the putative transition-state ensemble, the effect of macromolecular crowding on the relative stability between the unfolded ensemble and the transition-state ensemble was calculated. The resulting effect of crowding on the folding kinetics agrees well with experimental observations. This study shows that molecular dynamics simulations combined with calculation of crowding effects provide an avenue for characterize the transition-state ensemble in atomic details.

INTRODUCTION

Many single-domain proteins fold without detectable intermediates in kinetic folding experiments. However, intermediate states that are more stable than the unfolded state may still exist after crossing the rate-limiting transition state of folding (1,2). In fact, such hidden intermediates have been detected in several small proteins by using the native-state hydrogen exchange method (3–12). One of these proteins is Rd-apocyt b_{562} , a redesigned variant of the four-helix bundle protein apocytochrome b_{562} (13) (Fig. 1). Experimental studies on this protein have led to the identification of two partially unfolded forms (PUFs) after the rate-limiting transition state (8,14–17). PUF1 has the helix I and the C-terminal half of helix IV (or IV-C) unfolded but all other helical segments formed, whereas PUF2 has only helix I unfolded. A crude model for the folding pathway is thus (8,16)



where U, TS, and N denote the unfolded state, the transition state, and the native state, respectively; a pair of braces indicates a fast transition. NMR structures of mimetic models of the two PUFs have been solved (15,17). More recently PEG 20K has been used as a crowding agent to study the effect of macromolecular crowding on the folding kinetics (18). This multitude of experimental studies makes Rd-apocyt b_{562} a good model system for using molecular dynamics (MD) simulations to gain atomistic insight into the transition-state ensemble.

MD simulations have been used widely to characterize the transition-state ensembles in protein folding (19–27) and unfolding (28–45). In the latter simulations, high-temperature-induced unfolding has become a robust procedure since the pioneering work of Li and Daggett (28). Despite the successes and ongoing improvements of folding simulations, a number of advantages of unfolding simulations can be noted. Starting with the folded structure, a protein can be unfolded in a high-temperature simulation much faster than a simulation of the reverse process. Local-minimum traps are, by design, much easier to avoid in a high-temperature simulation. Folding simulations up to several microseconds are typically needed to fold small proteins (21–24,45), whereas simulations up to tens of nanoseconds are sufficient to unfold even medium-sized proteins. Beside the sampling problem, the inaccuracy of force fields is also a major obstacle to folding proteins correctly (46); a 10- μ s simulation of a small ultrafast-folding protein in explicit solvent failed to reach any structure resembling the native state (27). The principle of microscopic reversibility asserts that the pathways of folding and unfolding are the reverse of each other; hence the information from unfolding simulations can be used to deduce the folding pathway. Because the unfolding simulations are carried out at high temperatures, an added assumption is that unfolding pathways are robust so as to remain valid at the lower temperatures accessible by experiments. The question of whether unfolding pathways depend on temperature has been explored to some extent (34,40); these studies suggest that unfolding pathways are generally independent of temperature. This assumption is made here.

In this study, we carried out a set of high-temperature unfolding simulations of Rd-apocyt b_{562} and analyzed the

Submitted November 11, 2009, and accepted for publication January 5, 2010.

*Correspondence: hzhou4@fsu.edu

Hariato Tjong's present address is Department of Biological Sciences, University of Southern California, Los Angeles, CA.

Editor: Nathan Andrew Baker.

© 2010 by the Biophysical Society
0006-3495/10/05/2273/8 \$2.00

doi: 10.1016/j.bpj.2010.01.052

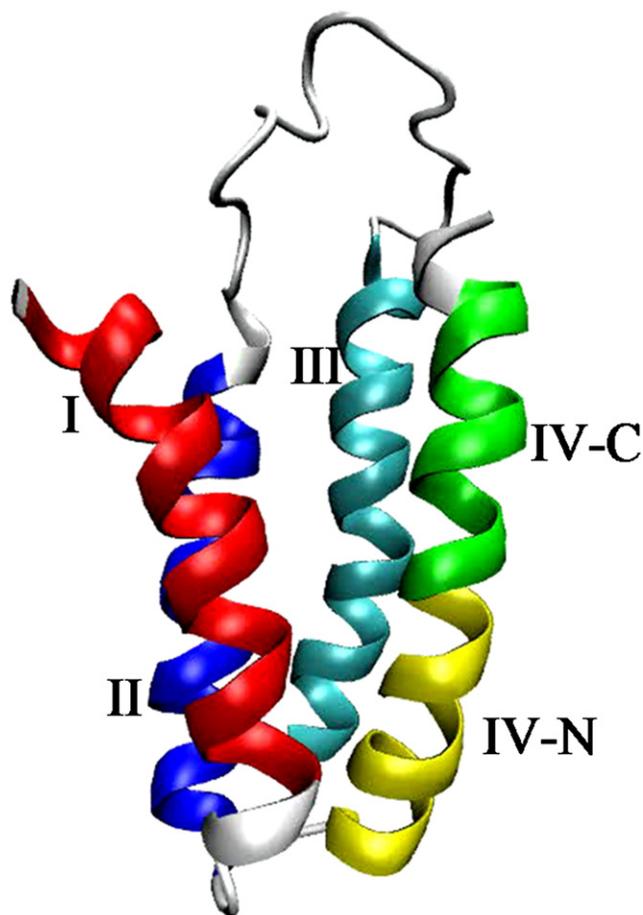


FIGURE 1 Structure of the four-helix bundle protein Rd-apocyct b_{562} . In the color version of this figure, helices I, II, III, IV-N, and IV-C are displayed in red, blue, cyan, yellow, and green, respectively.

unfolding pathway and the transition-state ensemble. The application of high-temperature unfolding simulations to Rd-apocyct b_{562} is particularly suitable because the PUFs identified by the experiments exist on the native side of the rate-limiting transition state, and hence are expected to occur early during unfolding. Indeed, as unfolding progressed in our simulations, helix I and then IV-C melted, producing PUF2 and PUF1, respectively. This unfolding order, on reversal, yields the same folding pathway as identified by the experimental studies (8,16). Furthermore, the higher stability of helices II, III, and IV-N (i.e., the N-terminal half of helix IV) over that of helices I and IV-C is correlated with the intrinsic helix propensities of the corresponding sequence segments.

Experimentally, Φ -value analysis yields valuable information about the transition-state ensemble (47). When complemented with MD simulations, an atomistic picture of the transition-state ensemble may be constructed (28–31,33,38,39). Here, based on the Φ -values (8,16), we empirically defined the transition-state ensemble in our simulations as comprised of all conformations in which helices I and IV-C have <3 residues remaining helical and at the same time

helices II, III, and IV-N have at least 50% of their residues remaining helical. With this definition, the residue helical probabilities of the transition-state ensemble show good correlation with the Φ -values.

To further validate the putative transition-state ensemble, we calculated how its stability is affected by crowding. The resulting effects on the folding kinetics are consistent with the measurements with PEG 20K as a crowding agent by Ai et al. (18). The consistency confirms our placement of the transition state.

SIMULATION, CALCULATION, AND METHODS

Unfolding and native-state simulations

All MD simulations of Rd-apocyct b_{562} started from the NMR structure (15) (PDB entry 1yyj). In the native state, the helical segments of this 106-residue protein are: helix I, residues 2–18; helix II, residues 23–40; helix III, residues 59–81, helix IV-N, residues 84–94; and helix IV-C, residues 95–104 (Fig. 1). The simulations followed the procedure in our previous study (40). Six unfolding simulations were carried out at 500 K using the SANDER module in AMBER package (48) with ff99SB force field (20,49,50). These simulations are referred to as MD1 to MD6. To explore conformational sampling in the native state, a simulation was also carried out at 300 K. In each simulation, the protein was solvated in 4918 TIP3P water molecules (51) and neutralized by four sodium ions. Each simulation was prepared by energy minimization, with a harmonic restraint on the protein atoms gradually decreasing from 50 kcal/mol/Å² to 0. For the native-state simulation, the energy-minimized structure was heated from 100 K to 300 K over 40 ps at constant volume. The simulation then continued at constant temperature and constant pressure to a total of 31 ns. For the high-temperature simulations, the energy-minimized structure was heated from 100 K to 500 K over 150 ps at constant volume, using different random number seeds. The simulations then continued at constant temperature and constant volume for a total of 15 ns each for MD1 to MD5 but for a total of 16 ns for MD6; the additional 1 ns of simulation is used to represent the unfolded ensemble (see below). The SHAKE algorithm (52) was used to constraint bonds involving hydrogen atoms. The timesteps were 2 fs for the 300-K simulation and 1 fs for the 500-K simulations. A cutoff of 10 Å for nonbonded interactions was used, and the particle mesh Ewald method (53–55) was used to treat the long range electrostatic interactions. After equilibration, the mass density was ~ 1.0 g/cm³ and 0.76 g/cm³ in the 300-K and 500-K simulations, respectively. Protein conformations were sampled every 1 ps for analysis.

Specification of transition-state, unfolded, and native ensembles

The Φ -values indicate that helices I and IV-C are unfolded and helices II, III, and IV-N are partially formed in the rate-limiting transition state (8,16). Accordingly we specified the transition-state ensemble in our simulations as comprised of all conformations in which helices I and IV-C have less than three residues remaining helical whereas helices II, III, and IV-N have at least 50% of their residues remaining helical. Specifically, helices II, III, and IV-N must have at least 9, 12, and 6 residues, respectively, still helical. This empirical definition yielded a total of 256 conformations from the six unfolding simulations. The secondary structures of individual residues in the protein conformations were assigned according to the DSSP program (56). A residue was classified as helical if the DSSP output was H and nonhelical otherwise. In the DSSP program, a helix is considered to be comprised of at least three consecutive residues. Our intention of limiting the number of helical residues to less than three for helices I and IV-C was to make these two helices completely unfolded in the transition-state ensemble. Either helix could still have a small helical content (as in fact is the case) because the first or last one or two residues of the helix

can be part of a longer helical stretch that goes beyond the N or C-terminal of the helix. Of course the residues beyond the N or C-terminal of the helix are not helical in the native state.

The unfolded and native ensembles were already used in our previous study on the effects of crowding on the folding stability of Rd-apocyt b_{562} (57). The unfolded ensemble consisted of the 1000 conformations from the additional 1 ns of MD6 after 15 ns of simulation. The native ensemble consisted of 1000 conformations sampled at 10 ps intervals from 6 to 16 ns of the native-state simulation.

Calculation of crowding effects

We have developed a postprocessing approach for simulating crowding (57,58). In the conventional approach of direct simulations (59–63), the test protein and a number of crowders are mixed and their motions are followed simultaneously. In our approach, we simulate the motion of the test protein in a particular state (e.g., the native state or the unfolded state), and the effects of crowding on the free energy of this state are predicted by the transfer free energy when representative protein conformations are placed inside a box of crowders. The approach is similar in spirit to Widom's particle insertion method (64). The postprocessing approach holds a number of advantages over direct simulations, leading to gain in calculation speed and accuracy (58). In particular, in all direct simulations, test proteins have always been represented at a coarse-grained level. In contrast, in our approach the simulation of the test protein is done in the absence of crowders and is restricted to a particular state. Therefore we can afford to represent the protein at an atomic level and exhaustively sample the state of interest.

The postprocessing approach has been applied to the native state and unfolded state of Rd-apocyt b_{562} to predict the effects of crowding on the folding stability (57). Here we applied it to the transition-state ensemble as well, yielding predictions for the effects of crowding on the folding kinetics. As before (57), we calculated crowding effects for a range of crowder sizes and a range of crowder concentrations. Specifically, crowder radii were 15, 20, and 30 Å, and crowder volume fractions were 5%, 15%, 25%, and 35%. The crowding-induced change in chemical potential for each of the 256 conformations in the transition-state ensemble was Boltzmann-averaged to yield the change in the free energy of the transition-state ensemble by crowding. This change was finally combined with the counterparts of the unfolded state and native state to predict the effects of crowding on the folding and unfolding activation free energies.

RESULTS AND DISCUSSION

The unfolding process

Fig. 2 displays the root mean-square deviations (RMSD) of the six high-temperature simulations and the native-state simulation from the starting NMR structure. The RMSD of the native-state simulation becomes steady in <1 ns and stays at relatively low values (<4 Å). In contrast, the RMSD of the high-temperature simulations show continued increase till the end of the 15 ns of simulation, reaching beyond 10 Å. At that point, the protein is mostly denatured (see below).

To gain further insight into the unfolding process, we used the DSSP program to calculate the secondary structures of individual residues along the six unfolding trajectories. The helical status of the residues as a function of the simulation time is displayed in Fig. 3. It is clear that, in all the six trajectories but MD1, the helices are melted by the end of the 15-ns simulations. Overall, helices I and IV-C melt relatively quickly, and helices II, III, and IV-N persist much longer

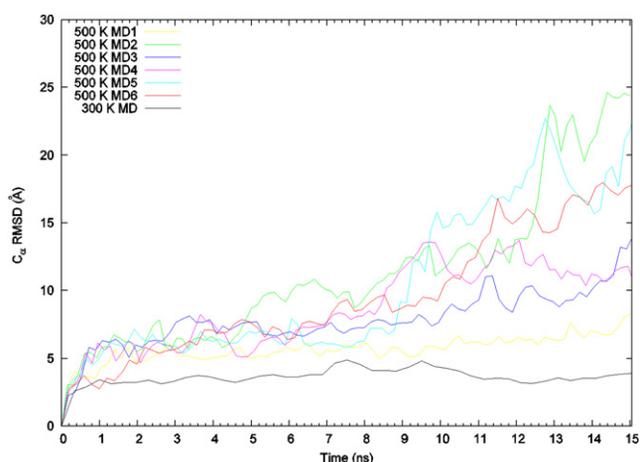


FIGURE 2 Root mean-square deviations of the simulations from the starting NMR structure, calculated over C_{α} atoms. The curves are smoothed using the bezier option in gnuplot. The RMSD of the native-state simulation remains at its low values all the way to the end of the 31-ns trajectory.

before melting. In MD1, first helices I and IV-C and then III and IV-N melt, but helix II remains mostly intact.

To be quantitative about the timing of helix melting, we calculated the mean first melting time. In each simulation, the first melting time of a particular helix was defined as the first time when 70% of its residues become melted. To prevent the calculation from being biased by the briefest excursions below the 70% threshold, the calculation of the first melting time was done after replacing the helical content at each time point by the running average over a 51-ps window. The calculation was then averaged over MD2 through MD6 (MD1 was excluded because helix II remains not completely melted by the end of the 15-ns simulation). The final results for the mean first melting times of helices I, II, III, IV-N, and IV-C are 1.4 ± 0.3 , 7.9 ± 3.7 , 5.7 ± 2.0 , 7.2 ± 3.0 , and 1.5 ± 1.1 ns. The conclusion is again that helices I and IV-C melt early, and after an extended period helices III, IV-N and II then start to melt in that order. Our simulations seem to suggest that there are two distinct unfolding events, one is the melting of the apparently weaker helices I and IV-C, the other is the melting of the apparently stronger helices II, III, and IV-C. The conformations sampled between these two events can be identified as constituting PUF1 (see below).

Unfolding pathway and helix propensities

To further quantify the strengths of the helices to withstand high temperature, we calculated the fraction of time that each residue is in the helical state during the 15-ns simulation. The profile of the helical fractions of the individual residues captures the relative stability of the different helices. Similar profiles are obtained from the separate simulations, but the magnitudes are somewhat different, reflecting the different paces of unfolding of the separate simulations. We scaled

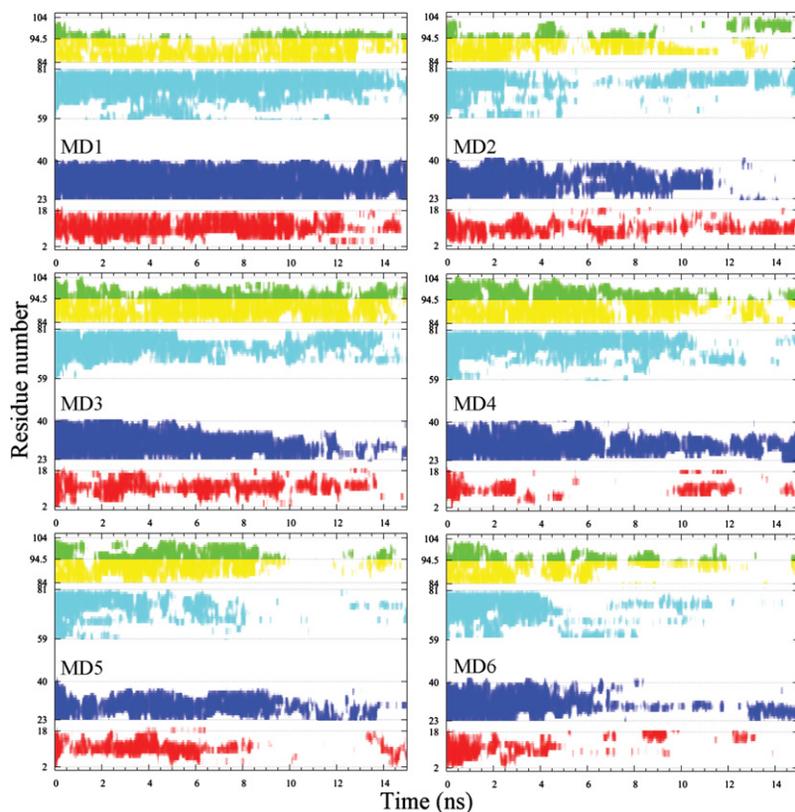


FIGURE 3 Helical status of the individual residues along the six high-temperature trajectories. Each helical residue is indicated by a vertical bar; the bars of the different helical segments have the same color code as in Fig. 1.

these profiles to the same maximum and averaged the results over MD2 through MD6 (again MD1 was excluded because of the incomplete melting of helix II). The composite profile of helical fractions is shown in Fig. 4. The peaks of helical fractions within the helices follow the order $I < IV-C < III < IV-N < II$. Not surprisingly, this order is the same as that of the mean first melting times.

Based on the above analysis, we may conclude that the unfolding pathway in our simulations is

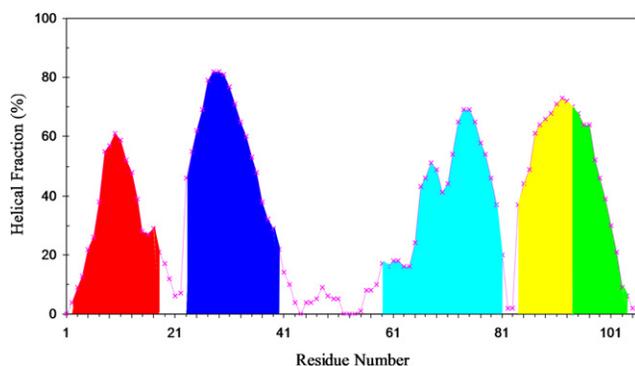
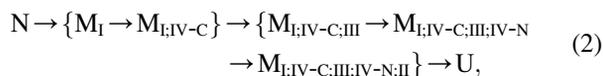


FIGURE 4 Composite profile of helical fractions calculated over the course of five high-temperature simulations. The areas above the individual helices of the native structure are separately shaded.

where M_I denotes intermediate state in which helix I is melted and other helices are still intact; the symbols for the other intermediates can be similarly interpreted. On reversal, this pathway corresponds well with folding scheme 1, proposed by Chu et al. (8) and Zhou et al. (16) based on their experimental studies. In particular, M_I and $M_{I,IV-C}$ correspond to PUF2 and PUF1, respectively. The transition state is passed when the other three helices are melted. We will return to the specification of the transition-state ensemble below.

What determines the relative stability of the helices to withstand high temperature and hence their order of melting? Our first instinct was that the intrinsic helix propensities dictated by the sequences play a role. To test this idea, we used the AGADIR program (<http://agadir.crg.es/>) to predict the helical contents from the corresponding sequences of the individual helical segments. The predictions are displayed in Fig. 5. They show surprisingly good match with the profile in Fig. 4. In particular, helices I and IV-C show lower peaks than helices II, III, and IV-C. Within each helix, the peaks occur at roughly the same residue positions in the two figures. For helix III, both figures show lower helical content in the N-terminal half than in the C-terminal half. Therefore intrinsic helix propensities seem to be a major determinant of the order of helix melting during the unfolding simulations. We caution, however, that a direct read-out of folding signals from protein sequence is challenging and no simple answer is expected. In this context, it is noted

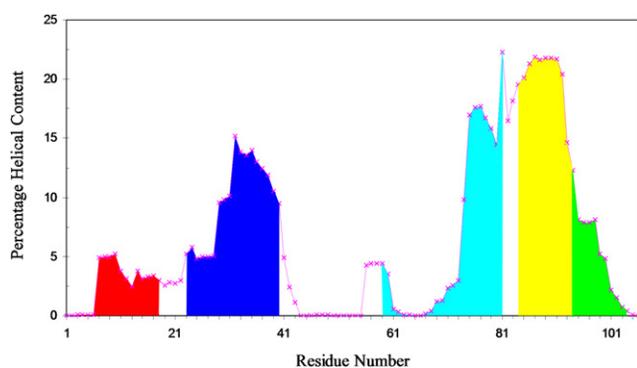


FIGURE 5 Intrinsic helix propensities predicted from the Rd-apocyto b_{562} sequence by the AGADIR program. Parameters used in the predictions were: 274 K; 0 ion strength; and pH 7. The areas above the individual helices of the native structure are separately shaded.

that the helix propensities of the sequence segments are all relatively low ($<25\%$), and potentially can be overridden by other factors such as tertiary interactions (see below).

The initiation of the unfolding process is of interest. In most of the simulations, the hydrophobic core formed by W7, P45, P46, I98, Y101, and Y105 is broken early. The side chain of Y101 rotates outward, toward helix III, followed by a similar rotation of the side chain of Y105. These rotations break their interactions with W7, P45, and P46. Soon after, the side chain of W7 interacts with that of I98 for sometime, whereas the N-terminal half of helix I and the C-terminal half of helix IV become frayed, further exposing the hydrophobic core to the solvent. The unfolding then proceeds to form the M_I and $M_{I,IV-C}$ intermediates.

It is worth noting that Feng et al. (10) have detected an additional state, denoted by N^* , in equilibrium unfolding studies using urea denaturation. The N^* state has the native fold but nonnative hydrophobic interactions. Along the folding pathway, N^* is placed between PUF2 and N. The main difference between N and N^* is in hydrophobic core packing: in N, Y101 and Y105 interact with W7; but in N^* , Y101 and Y105 face outward and become closer to helix III, and W7 interacts with I98. These signatures of N^* and its placement between N and PUF2 are just as observed during the initiation of the unfolding process, as outlined in the preceding paragraph. Feng et al. (15) have further detected nonnative hydrophobic interactions in PUF2. We observe very similar nonnative interactions in our M_I intermediate, e.g., between F65 and F69, and between L68 and I98.

Transition-state ensemble

As described above, the unfolding of Rd-apocyto b_{562} consists of two distinct events: the early melting of helices I and IV-C and the melting of the remaining helices after a relatively long delay. Kinetically it seems logical to associate the crossing of the rate-limiting transition state with the second event. (This association is consistent with a suggestion

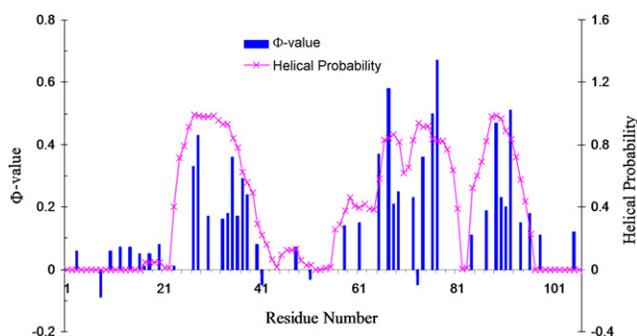


FIGURE 6 Comparison residue helix probabilities of the transition-state ensemble extracted from the simulations and Φ -values (8,16).

made in our previous study using high-temperature simulations to characterize folding and unfolding pathways (40): the last remaining secondary structures should be part of the transition state for folding.) The Φ -values also indicate that helices I and IV-C are unfolded and helices II, III, and IV-N are partially formed in the transition state (8,16) (Fig. 6). These observations led us to define the transition-state ensemble as comprised of all conformations in which helices I and IV-C have less than three helical residues whereas helices II, III, and IV-N have at least 50% of helical residues. A total of 256 such conformations are collected from the six unfolding simulations. On average, the percentages of helical residues in the transition-state ensemble are $(0.6 \pm 2.5)\%$, $(78.6 \pm 16.0)\%$, $(67.5 \pm 13.0)\%$, $(78.1 \pm 16.8)\%$, and $(6.6 \pm 8.2)\%$ for helices I, II, III, IV-N, and IV-C, respectively. The significant variances of the percent helicities give an indication of the diversity of the transition-state ensemble.

To assess this putative transition-state ensemble, we calculated the helical probabilities of individual residues (i.e., the percentage of times that a given residue is in the helical state in the 256-conformation ensemble). In Fig. 6 we compare the helical probabilities with the experimental Φ -values. A strong correlation is observed. As expected, the very low Φ -values of helices I and IV-C and the relatively high Φ -values of helices II, III, and IV-N are mapped by the corresponding disparities in helical probability among the different helices. Furthermore, the peak positions of Φ -values within helices II, III, and IV-N are well reproduced by the helical probabilities. Note that the correlation between the helical probabilities and the Φ -values is not intended as showing agreement between calculation and experiment; rather, it is meant to confirm that our putative transition-state ensemble, with atomic details not accessible to experiments, is quantitatively validated by Φ -value analysis.

To find a conformation that may serve as a representative of the transition-state ensemble, we carried out a cluster analysis according to RMSD of C_α atoms (65). Using a 8-Å RMSD cutoff, the 256 conformations are found to form a major cluster, including 207 conformations (81% of the total), and several much smaller clusters. The representative

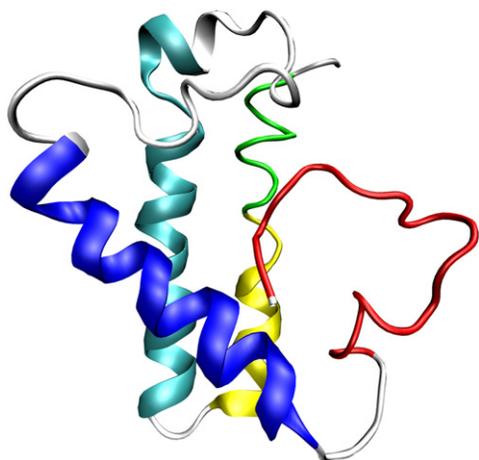


FIGURE 7 Structure of a representative of the transition-state ensemble.

of this cluster is shown in Fig. 7. In this structure, helices II and III are still mostly intact, but they are no longer parallel as they are in the native state (see Fig. 1). As a result, these helices lose some of their native interactions. Our structural interpretation of the experimental observation that even the highest Φ -values, occurring in helices II, III, and IV-C, are significantly less than one is that it does not necessarily mean these helices are partially or largely unfolded in the transition-state ensemble. Rather, these helices may still be mostly intact, but their relative orientations are altered, thereby losing some of their native interactions, as illustrated by the representative transition-state structure of Fig. 7.

From the transition-state ensemble, further unfolding is fast. Additional native interactions are lost as helices II and III further change their relative orientation and separation. Thereafter helix III starts to melt from the N-terminal half.

Macromolecular crowding

Although our proposed unfolding pathway and transition-state ensemble seem to rationalize a host of experimental observations, it remains important to find independent

validation of the putative transition-state ensemble. To that end, we calculated the effects of macromolecular crowding on the free energy of the transition-state ensemble, following our previous work presenting similar calculations on the unfolded and native states (57).

In Fig. 8, we display the predicted effects of crowding on the folding and unfolding activation free energies, for a range of crowder radii (15–30 Å) and a range of crowder volume fractions (5–35%). Crowding is seen to exert modest stabilization of the transition-state ensemble relative to the unfolded state, and vanishingly small destabilization of the transition-state ensemble relative to the native state. These results are in qualitative agreement with the measurements by Ai et al. (18), who found that the addition of 85 g/l of PEG 20K results in a modest increase in the folding rate of Rd-apocyt b_{562} but essentially no effect on the unfolding rate. Specifically, an increase of 45% in folding rate and a decrease of 6% in unfolding rate are observed at 25°C. Such effects are predicted by our calculation for crowders with a 20-Å and occupying 15% of volume. Atomistic representation of the crowding agent in future work will make the comparison more quantitative.

In our calculation, the crowding-induced decrease in folding activation free energy originates from the fact that the transition-state conformations are more compact than the unfolded conformations. In general, compact conformations are less destabilized by crowding than open conformations. The modest increase in folding rate by the addition of PEG 20K is thus an indication that the conformations of Rd-apocyt b_{562} experience a compaction when it moves from the unfolded state to the rate-limiting transition state. Similarly, the negligible effect of PEG 20K on the folding rate is an indication that the transition state is not significantly less compact than the native state. The experimental results on crowding effects thus provide both qualitative and quantitative support of our putative transition-state ensemble. Use of different crowding agents and their mixtures (66) will further increase the discriminatory power of macromolecular crowding as a probe of transition-state ensembles.

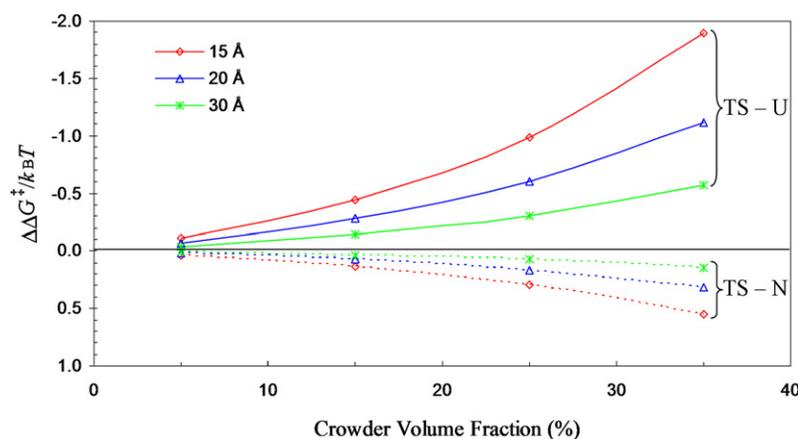


FIGURE 8 Effects of crowding on the folding and unfolding activation free energies, $\Delta\Delta G^\ddagger_{TS-U}$ and $\Delta\Delta G^\ddagger_{TS-N}$, in units of $k_B T$ where k_B is Boltzmann constant and T is the absolute temperature. Results for three crowder radii (with values shown) are displayed.

CONCLUSIONS

By high-temperature simulations, we observed the following unfolding pathway of Rd-apocyt b_{562} : helices I and IV-C melt quickly, and after an extended delay, helices III, IV-N, and II start to melt; the melting of the latter helices finally unfold the protein completely. The reversal of the unfolding pathway can be proposed as the folding pathway. It seems that the folding/unfolding pathway is governed mainly by intrinsic helix propensity, whereby helices with higher helix propensities form first in the folding process. Our simulation results reproduce well experimental observations (8,16,17) and provide a rationalization. In addition to the folding/unfolding pathway, the simulations yield an atomic representation of the transition-state ensemble. The putative transition-state ensemble can be further validated, as we do here by calculation of crowding effects. We show that a modest increase in folding rate and null effect on the unfolding rate by the addition of a crowding agent (18) both qualitatively and quantitatively support the transition-state ensemble extracted from the simulations. MD simulations combined with calculation of crowding effects thus provide an avenue for characterizing the transition-state ensemble in atomic details.

This work was supported in part by the National Institutes of Health (grant No. GM058187).

REFERENCES

- Bai, Y., and S. W. Englander. 1996. Future directions in folding: the multi-state nature of protein structure. *Proteins*. 24:145–151.
- Bai, Y. 2006. Energy barriers, cooperativity, and hidden intermediates in the folding of small proteins. *Biochem. Biophys. Res. Commun.* 340:976–983.
- Bai, Y., T. R. Sosnick, ..., S. W. Englander. 1995. Protein folding intermediates: native-state hydrogen exchange. *Science*. 269:192–197.
- Chamberlain, A. K., T. M. Handel, and S. Marqusee. 1996. Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH. *Nat. Struct. Mol. Biol.* 3:782–787.
- Fuentes, E. J., and A. J. Wand. 1998. Local stability and dynamics of apocytochrome b_{562} examined by the dependence of hydrogen exchange on hydrostatic pressure. *Biochemistry*. 37:9877–9883.
- Llinas, M., B. Gillespie, ..., S. Marqusee. 1999. The energetics of T4 lysozyme reveal a hierarchy of conformations. *Nat. Struct. Mol. Biol.* 6:1072–1078.
- Fersht, A. R. 2000. A kinetically significant intermediate in the folding of barnase. *Proc. Natl. Acad. Sci. USA*. 97:14121–14126.
- Chu, R., W. Pei, ..., Y. Bai. 2002. Relationship between the native-state hydrogen exchange and folding pathways of a four-helix bundle protein. *Biochemistry*. 41:7998–8003.
- Yan, S., S. D. Kennedy, and S. Koide. 2002. Thermodynamic and kinetic exploration of the energy landscape of *Borrelia burgdorferi* OspA by native-state hydrogen exchange. *J. Mol. Biol.* 323:363–375.
- Feng, H., N.-D. Vu, and Y. Bai. 2004. Detection and structure determination of an equilibrium unfolding intermediate of Rd-apocytochrome b_{562} : native fold with non-native hydrophobic interactions. *J. Mol. Biol.* 343:1477–1485.
- Vu, N.-D., H. Feng, and Y. Bai. 2004. The folding pathway of barnase: the rate-limiting transition state and a hidden intermediate under native conditions. *Biochemistry*. 43:3346–3356.
- Feng, H., N.-D. Vu, and Y. Bai. 2005. Detection of a hidden folding intermediate of the third domain of PDZ. *J. Mol. Biol.* 346:345–353.
- Feng, H., and Y. Bai. 2004. Repacking of hydrophobic residues in a stable mutant of apocytochrome b_{562} selected by phage-display and proteolysis. *Proteins*. 56:426–429.
- Takei, J., W. Pei, ..., Y. Bai. 2002. Populating partially unfolded forms by hydrogen exchange-directed protein engineering. *Biochemistry*. 41:12308–12312.
- Feng, H., J. Takei, ..., Y. Bai. 2003. Specific non-native hydrophobic interactions in a hidden folding intermediate: implications for protein folding. *Biochemistry*. 42:12461–12465.
- Zhou, Z., Y. Huang, and Y. Bai. 2005. An on-pathway hidden intermediate and the early rate-limiting transition state of Rd-apocytochrome b_{562} characterized by protein engineering. *J. Mol. Biol.* 352:757–764.
- Feng, H., Z. Zhou, and Y. Bai. 2005. A protein folding pathway with multiple folding intermediates at atomic resolution. *Proc. Natl. Acad. Sci. USA*. 102:5026–5031.
- Ai, X., Z. Zhou, ..., W. Y. Choy. 2006. ^{15}N NMR spin relaxation dispersion study of the molecular crowding effects on protein folding under native conditions. *J. Am. Chem. Soc.* 128:3916–3917.
- Snow, C. D., H. Nguyen, ..., M. Gruebele. 2002. Absolute comparison of simulated and experimental protein-folding dynamics. *Nature*. 420:102–106.
- Simmerling, C., B. Strockbine, and A. E. Roitberg. 2002. All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.* 124:11258–11259.
- Chowdhury, S., M. C. Lee, ..., Y. Duan. 2003. Ab initio folding simulation of the Trp-cage mini-protein approaches NMR resolution. *J. Mol. Biol.* 327:711–717.
- Pitera, J. W., and W. Swope. 2003. Understanding folding and design: replica-exchange simulations of “Trp-cage” miniproteins. *Proc. Natl. Acad. Sci. USA*. 100:7587–7592.
- Lei, H., and Y. Duan. 2007. Two-stage folding of HP-35 from ab initio simulations. *J. Mol. Biol.* 370:196–206.
- Yang, L., M. P. Grubb, and Y. Q. Gao. 2007. Application of the accelerated molecular dynamics simulations to the folding of a small protein. *J. Chem. Phys.* 126:125102–125107.
- Ensign, D. L., P. M. Kasson, and V. S. Pande. 2007. Heterogeneity even at the speed limit of folding: large-scale molecular dynamics study of a fast-folding variant of the villin headpiece. *J. Mol. Biol.* 374:806–816.
- Maragakis, P., K. Lindorff-Larsen, ..., D. E. Shaw. 2008. Microsecond molecular dynamics simulation shows effect of slow loop dynamics on backbone amide order parameters of proteins. *J. Phys. Chem. B*. 112:6155–6158.
- Freddolino, P. L., F. Liu, ..., K. Schulten. 2008. Ten-microsecond molecular dynamics simulation of a fast-folding WW domain. *Biophys. J.* 94:L75–L77.
- Li, A., and V. Daggett. 1994. Characterization of the transition state of protein unfolding by use of molecular dynamics: chymotrypsin inhibitor 2. *Proc. Natl. Acad. Sci. USA*. 91:10430–10434.
- Li, A., and V. Daggett. 1996. Identification and characterization of the unfolding transition state of chymotrypsin inhibitor 2 by molecular dynamics simulations. *J. Mol. Biol.* 257:412–429.
- Lazaridis, T., and M. Karplus. 1997. “New view” of protein folding reconciled with the old through multiple unfolding simulations. *Science*. 278:1928–1931.
- Li, A., and V. Daggett. 1998. Molecular dynamics simulation of the unfolding of barnase: characterization of the major intermediate. *J. Mol. Biol.* 275:677–694.
- Tsai, J., M. Levitt, and D. Baker. 1999. Hierarchy of structure loss in MD simulations of src SH3 domain unfolding. *J. Mol. Biol.* 291:215–225.
- Gsponer, J., and A. Caffisch. 2001. Role of native topology investigated by multiple unfolding simulations of four SH3 domains. *J. Mol. Biol.* 309:285–298.

34. Day, R., B. J. Bennion, ..., V. Daggett. 2002. Increasing temperature accelerates protein unfolding without changing the pathway of unfolding. *J. Mol. Biol.* 322:189–203.
35. Daggett, V. 2002. Molecular dynamics simulations of the protein unfolding/folding reaction. *Acc. Chem. Res.* 35:422–429.
36. Day, R., and V. Daggett. 2005. Ensemble versus single-molecule protein unfolding. *Proc. Natl. Acad. Sci. USA.* 102:13445–13450.
37. Snow, C. D., E. J. Sorin, ..., V. S. Pande. 2005. How well can simulation predict protein folding kinetics and thermodynamics? *Annu. Rev. Biophys. Biomol. Struct.* 34:43–69.
38. White, G. W. N., S. Gianni, ..., V. Daggett. 2005. Simulation and experiment conspire to reveal cryptic intermediates and a slide from the nucleation-condensation to framework mechanism of folding. *J. Mol. Biol.* 350:757–775.
39. Daggett, V. 2006. Protein folding-simulation. *Chem. Rev.* 106:1898–1916.
40. Huang, X., and H.-X. Zhou. 2006. Similarity and difference in the unfolding of thermophilic and mesophilic cold shock proteins studied by molecular dynamics simulations. *Biophys. J.* 91:2451–2463.
41. Eleftheriou, M., R. S. Germain, ..., R. Zhou. 2006. Thermal denaturing of mutant lysozyme with both the OPLSAA and the CHARMM force fields. *J. Am. Chem. Soc.* 128:13388–13395.
42. Das, A., and C. Mukhopadhyay. 2007. Application of principal component analysis in protein unfolding: an all-atom molecular dynamics simulation study. *J. Chem. Phys.* 127:165103.
43. Beck, D. A. C., and V. Daggett. 2007. A one-dimensional reaction coordinate for identification of transition states from explicit solvent $P_{(fold)}$ -like calculations. *Biophys. J.* 93:3382–3391.
44. Settanni, G., and A. R. Fersht. 2008. High temperature unfolding simulations of the TRPZ1 peptide. *Biophys. J.* 94:4444–4453.
45. Yang, L., Q. Shao, and Y. Q. Gao. 2009. Thermodynamics and folding pathways of trpzip2: an accelerated molecular dynamics simulation study. *J. Phys. Chem. B.* 113:803–808.
46. Freddolino, P. L., S. Park, ..., K. Schulten. 2009. Force field bias in protein folding simulations. *Biophys. J.* 96:3772–3780.
47. Fersht, A. R., A. Matouschek, and L. Serrano. 1992. The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J. Mol. Biol.* 224:771–782.
48. Case, D. A., T. A. Darden, ..., P. A. Kollman. 2004. AMBER 8. University of California, San Francisco, CA.
49. Wang, J. M., P. Cieplak, and P. A. Kollman. 2000. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* 21:1049–1074.
50. Hornak, V., R. Abel, ..., C. Simmerling. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins.* 65:712–725.
51. Jorgensen, W. L., J. Chandrasekhar, ..., M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
52. Ryckaert, J.-P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23:327–341.
53. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.
54. Essmann, U., L. Perera, ..., L. G. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.
55. Saguí, C., and T. A. Darden. 1999. Molecular dynamics simulations of biomolecules: long-range electrostatic effects. *Annu. Rev. Biophys. Biomol. Struct.* 28:155–179.
56. Kabsch, W., and C. Sander. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22:2577–2637.
57. Qin, S., and H.-X. Zhou. 2010. Atomistic modeling of macromolecular crowding predicts modest increases in protein folding and binding stability. *Biophys. J.* 97:12–19.
58. Qin, S., D. D. Minh, ..., H.-X. Zhou. 2009. A method to predict crowding effects by postprocessing molecular dynamics trajectories: application to the flap dynamics of HIV-1 protease. *J. Phys. Chem. Lett.* 1:107–110.
59. Cheung, M. S., D. Klimov, and D. Thirumalai. 2005. Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc. Natl. Acad. Sci. USA.* 102:4753–4758.
60. Stagg, L., S. Q. Zhang, ..., P. Wittung-Stafshede. 2007. Molecular crowding enhances native structure and stability of alpha/beta protein flavodoxin. *Proc. Natl. Acad. Sci. USA.* 104:18976–18981.
61. Mittal, J., and R. B. Best. 2009. Theory and simulation of macromolecular crowding effects on protein folding stability and kinetics. *Biophys. J.* Accepted.
62. Pincus, D. L., and D. Thirumalai. 2009. Crowding effects on the mechanical stability and unfolding pathways of ubiquitin. *J. Phys. Chem.* 113:359–368.
63. Minh, D. D., C. E. Chang, ..., J. A. McCammon. 2006. The influence of macromolecular crowding on HIV-1 protease internal dynamics. *J. Am. Chem. Soc.* 128:6006–6007.
64. Widom, B. 1963. Some topics in the theory of fluids. *J. Chem. Phys.* 39:2802–2812.
65. de Vries, S. J., A. D. J. van Dijk, ..., A. M. J. J. Bonvin. 2007. HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins.* 69:726–733.
66. Batra, J., K. Xu, and H. X. Zhou. 2009. Nonadditive effects of mixed crowding on protein stability. *Proteins.* 77:133–138.