Urea Impedes the Hydrophobic Collapse of Partially Unfolded Proteins

Martin C. Stumpe and Helmut Grubmüller*
Max-Planck-Institute for Biophysical Chemistry, Theoretical and Computational Biophysics Department, Göttingen, Germany

ABSTRACT Proteins are denatured in aqueous urea solution. The nature of the molecular driving forces has received substantial attention in the past, whereas the question how urea acts at different phases of unfolding is not yet well understood at the atomic level. In particular, it is unclear whether urea actively attacks folded proteins or instead stabilizes unfolded conformations. Here we investigated the effect of urea at different phases of unfolding by molecular dynamics simulations, and the behavior of partially unfolded states in both aqueous urea solution and in pure water was compared. Whereas the partially unfolded protein in water exhibited hydrophobic collapses as primary refolding events, it remained stable or even underwent further unfolding steps in aqueous urea solution. Further, initial unfolding steps of the folded protein were found not to be triggered by urea, but instead, stabilized. The underlying mechanism of this stabilization is a favorable interaction of urea with transiently exposed, less-polar residues and the protein backbone, thereby impeding back-reactions. Taken together, these results suggest that, quite generally, urea-induced protein unfolding proceeds primarily not by active attack. Rather, thermal fluctuations toward the unfolded state are stabilized and the hydrophobic collapse of partially unfolded proteins toward the native state is impeded. As a result, the equilibrium is shifted toward the unfolded state.

INTRODUCTION

Urea is a strong denaturant for proteins. However, despite its everyday use to study protein folding and stability, the molecular mechanism by which urea acts as denaturant is still not well understood. In particular, the question of which kind of interactions embody the main driving force for urea-induced denaturation has been intensively studied. Growing evidence has accumulated that preferential solvation of less polar protein parts by urea molecules weakens the hydrophobic effect and hence leads to protein denaturation (1–11), although alternative views also exist (12–15).

Here, we investigate by molecular dynamics simulations how urea affects protein conformations at different states of folding/unfolding. In particular, we address the question: does urea actively destabilize the folded state, or instead stabilize the denatured state? Calorimetric studies point toward a stabilization of the denatured state (16), but the underlying processes at the molecular level are unclear.

In recent simulation studies, the effect of urea on model systems for proteins or hydrophobic amino acids has been investigated. Particularly, it was found that urea inhibits dewetting of hydrophobic surfaces (11), and that urea stabilizes a pair of separated neopentanes in water (7). Both studies suggest a weakening of the hydrophobic effect by urea, with implications for the mechanism of how urea denatures proteins.

For proteins, however, simulation studies of the effect of urea are much more challenging. In particular, denaturation typically occurs on timescales orders-of-magnitude out of reach of current atomistic computer simulations. Indeed, simulations of the CI2 protein (9) have revealed only minor urea-induced perturbations of the native state even within several hundred nanoseconds. To overcome this problem, here a number of partially unfolded structures of the Cold Shock protein Bc-CsP from *Bacillus caldolyticus* are generated by a high-temperature unfolding simulation. Subsequently, the dynamics of these partially unfolded structures are compared in water and in urea at room temperature. This approach enables us to study the effect of urea on different states of folding/unfolding at atomic detail within feasible simulation times.

The interaction of urea and water with individual amino acids in tripeptides without tertiary structure has been extensively studied recently (8). In this study, the simulation of partially unfolded structures also allows us to quantify these interactions residuewise. By comparison, the effect of sequence/structure and solvent accessibility of amino acids within the protein scaffold will be deduced.

METHODS

Simulation setup

For all simulations, the molecular dynamics package GROMACS (17,18) program suite, Vers. 3.2.1 and 3.3, was used with the OPLS-all-atom force field (19). TIP4P (20) was used for water, and urea parameters were taken from Smith et al. (21). Particle-mesh Ewald summation (22,23) was used to calculate the long-range electrostatic interactions with a grid-spacing of 0.12 nm and an interpolation order of 4. A cutoff of 1.0 nm was used for the short-range Coulomb as well as the Lennard-Jones interactions. All simulations were performed with a 2-fs integration time step in the *NpT*-ensemble using Berendsen-type temperature-coupling (24) with a coupling coefficient of $\tau_T = 0.1$ ps and Berendsen-type pressure-coupling (24) at 1 bar with a coupling coefficient of $\tau_p = 1$ ps.

The crystal structure of the protein was taken from the Protein Data Bank (25), PDB code No. 1C9O (26). The size of the rectangular simulation box was chosen such that a minimum distance of 1.5 nm between protein atoms and the box was kept in each direction. For the solvation of the protein,

Submitted December 2, 2008, and accepted for publication January 27, 2009.

*Correspondence: hgrubmu@gwdg.de

Editor: Marilyn Gunner.

0006-3495/09/05/3744/9 $2.00

doi: 10.1016/j.bpj.2009.01.051

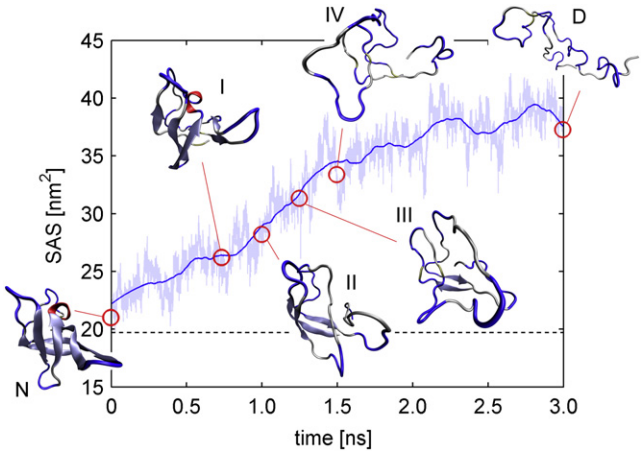

FIGURE 1 SAS of the protein in the initial high-temperature unfolding simulation. The bold line shows the running average over 250 ps, the dim line shows raw data. Also shown are snapshots of the four partially unfolded structures (I–IV) selected for subsequent room temperature simulations, together with the native state (*N*) and the selected completely unfolded structure (*D*). The dotted line denotes the SAS of the folded protein.

preequilibrated structures of water and 8 M urea (27) were used. Sodium and chloride ions were added to obtain a physiological ion concentration of 150 mM.

Each simulation was preceded by a 200-step steepest-descent energy minimization and a 500-ps solvent equilibration with position restraints on the protein heavy atoms. This equilibration protocol was also used for the partially unfolded structures from the high temperature simulation after replacing water with 8 M urea solution.

Generation of partially unfolded conformations

A simulation at 700 K in water was performed to quickly unfold the protein and to generate partially unfolded conformations. Constant volume condition (NVT ensemble) was used for this high-temperature simulation to prevent solvent evaporation, and the time step was set to 1 fs for numerical stability. Fig. 1 shows the solvent-accessible hydrophobic surface (SAS) for this heat-unfolding simulation as well as four partially unfolded structures (*I–IV*) together with the native structure (*N*) and one completely unfolded structure (*D*). The native structure (N), structures I–IV, and D were used as starting structures for the subsequent room temperature simulations in water or urea at 300 K, as listed in Table 1. To avoid overinterpretation of possibly anecdotal events, multiple simulation runs were carried out for the starting structures III and IV, on which our main conclusions are based. The total simulation time was $\approx$4.5 μs.

Analysis

SAS areas were calculated using the double cubic lattice method (28) with a 0.14-nm probe radius. Native contacts and native secondary structure were defined using the native state simulation in water, rather than the crystal structure. This approach has the advantage that fluctuations of the native state were captured that allowed a more direct comparison with the unfolding simulations. Residues were defined to be in contact if the distance between the closest atom pair was not larger than 0.4 nm. Contacts were defined as native if they were present during >50% of the time in the simulation of the native state in water. Contacts between neighboring residues were not considered for the calculation of the native contact fraction.

Secondary structure was classified using DSSP (29). The native secondary structure was defined as the most frequently occurring structure type for each residue seen in the simulation of the native state in water, which was similar to that of the crystal structure. Helix, β-sheet, and turn elements were considered to calculate the fraction of native secondary structure content.

TABLE 1 Summary of all 20 simulations performed for the Cold Shock protein, starting from the native structure (N), the partially unfolded structures (I–IV), or the completely unfolded structure (D)

Starting structure		Solvent	Simulation time
N	(0 ps)	water	219 ns
N	(0 ps)	8 M urea	453 ns
I	(750 ps)	water	209 ns
I	(750 ps)	8 M urea	151 ns
II	(1000 ps)	water	167 ns
II	(1000 ps)	8 M urea	143 ns
III	(1250 ps) × 3	water	139 ns/250 ns/244 ns
III	(1250 ps) × 3	8 M urea	119 ns/254 ns/255 ns
IV	(1500 ps) × 3	water	196 ns/138 ns/230 ns
IV	(1500 ps) × 3	8 M urea	142 ns/274 ns/276 ns
D	(3000 ps)	water	308 ns
D	(3000 ps)	8 M urea	314 ns

Contact coefficient

To quantify the frequency of interactions between urea and the amino acids, we used the contact coefficient C_{UW} (8) for a particular amino acid X,

$$C_{UW_X} = \frac{N_{X-U}}{N_{X-W}} \cdot \frac{M_W}{M_U}, \quad (1)$$

where N_{X-U} and N_{X-W} are the numbers of atomic contacts during the simulation of amino acid X with urea and water molecules, respectively. Atoms were defined to be in contact if they were closer than 0.35 nm. C_{UW} is normalized using the total numbers of urea atoms (M_U) and water atoms (M_W). Accordingly, a residue with a contact coefficient of $C_{UW} = 1.0$ has no interaction preference for either urea or water. Values >1.0 indicate preferential interaction with urea, values <1.0 indicate preferential interaction with water. The analysis was performed for a 300-ns trajectory of the unfolded state (starting with structure D) to minimize bias from geometric effects of the folded structure, and a 300-ns trajectory of the folded state (N) for comparison.

Estimation of free energy profiles from a Markov model

We estimated the free energy of the protein in solvation as a function of its SAS. The following procedure was used:

1. The SAS was discretized into N equidistant bins B_i ($i = 1 \ldots N$) between the minimum and the maximum SAS observed in our simulations.
2. N Markov states M_i were defined, where M_i is the ensemble subset of all protein conformations with an SAS within bin B_i.
3. For every time step, the protein was assigned to one Markov state M_i, according to the SAS of that conformation.
4. All transition times for individual jumps between adjacent Markov states (i.e., folding or unfolding steps) were extracted from the simulations.
5. For each transition, the underlying rate k was estimated from the set of corresponding transition times t_j ($j = 1 \ldots N_t$), where N_t is the number of observed transitions), using a maximum likelihood approach, where the probability p to observe the particular transition times given the rate k was maximized. Assuming that the transitions have a constant probability and are statistically independent from each other (Poisson process), this probability is

$$p(t_j|k) = \prod_{j=1}^{N_t} p(t_j|k) = \prod_{j=1}^{N_t} k e^{-kt_j}.$$

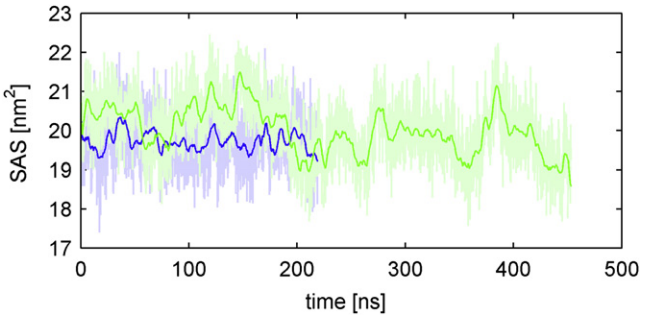

FIGURE 2 SAS of the protein at room temperature. (*Bold line*) Running average over 250 ps; (*thin line*) raw data, (*blue*) in water; (*green*) in 8 M urea.

As can easily be shown, it is maximal for $k^{-1} = \langle t \rangle$, i.e., when the inverse rate is the average of all observed transition times t_j ($j = 1 \ldots N_t$).

6. The free energy difference ΔG between two adjacent states M_i and M_{i+1} was calculated from the corresponding unfolding and folding transition rates (k_i^u and k_{i+1}^f, respectively) according to

$$\Delta G = G_{i+1} - G_i = -k_B T \ln \frac{k_i^u}{k_{i+1}^f}.$$

In addition to assuming a Poisson process for the transitions between states, this approach further assumes the states B_i ($i = 1 \ldots N$) to be Markovian, which is most likely not the case. In particular, since protein conformations are projected on a one-dimensional reaction coordinate here, each of the states B_i ($i = 1 \ldots N$) comprises, structurally, quite heterogenous conformations with boundaries between adjacent states that are not unambiguously defined; hence, transition probabilities will not be memory-free. This is particularly true for larger SAS values to which structurally very heterogenous conformations contribute. However, deviations from Markovian behavior were found to be <15%.

We note that we do not expect this approach to yield accurate free energy values, but rather to provide a rough estimate of qualitative features of the free energy landscape. Thus, it will mostly serve to illustrate and support results from other analyses, rather than be the basis for independent conclusions.

RESULTS AND DISCUSSION

Dynamics of the native state

The SAS area of the protein will be used as reaction coordinate for the transition from the folded state to the unfolded state. Fig. 2 shows the SAS for the simulations of the native state, starting from the crystal structure. As can be seen, the SAS of the protein does not increase significantly during the simulation with water (*blue line*) or with 8 M urea as solvent (*green line*). Thus, the folded state of the protein is stable within the simulation times, as is also confirmed by other measures for conformational changes, such as radius of gyration or root-mean-square-deviation of the backbone atoms (data not shown). As expected from the experimental timescale for denaturation of Cold Shock proteins (30,31), no unfolding is observed.

Nevertheless, as had been observed previously for the CI2 protein (9), the SAS of the Cold Shock protein, too, is larger in urea than in water, although the effect is less pronounced. Here, for the Cold Shock protein, this difference is much smaller on average ($\approx$0.3 nm^2) than for the CI2 protein ($\approx$2.0 nm^2). Further, the SAS fluctuations are large. Closer analysis reveals that residues Phe^{27} and Arg^{56} contribute significantly to this SAS difference (Fig. 3). The fluctuations in the total SAS result primarily from disruption and reformation of contact between these residues, accompanied by detachment and reattachment of the corresponding turn region. Whereas this contact between Phe^{27} and apolar parts of Arg^{56} is persistent in the simulation with water (Fig. 3 *a*), it is destabilized in urea (Fig. 3 *b*), where urea molecules disrupt the stacking interaction, as can be seen from the snapshots in Fig. 3 *c*. Most notably, this stacking contact ruptures several times in the simulation in water as well, but the contact reforms again after a few nanoseconds. Hence, this fluctuation is not triggered, but instead, stabilized by urea.

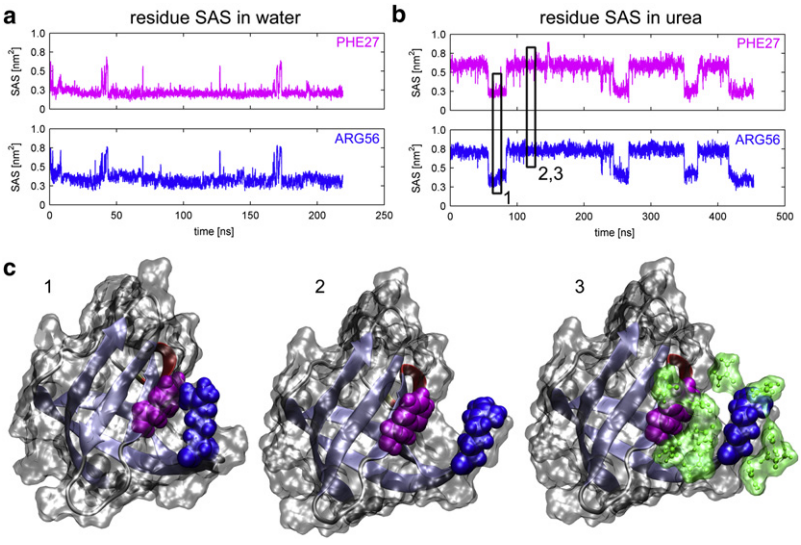

FIGURE 3 The contact between Phe^{27} and Arg^{56} is disrupted in urea, which gives rise to the fluctuations in the total SAS. (*a*) SAS of Phe^{27} and Arg^{56} in water, (*b*) SAS of Phe^{27} and Arg^{56} in urea, and (*c*) snapshots of the protein (snapshot 1), with Phe^{27} (*magenta*) and Arg^{56} (*blue*) in contact (snapshot 2), with Phe^{27} and Arg^{56} not in contact (snapshot 3), same as in snapshot 2 with adjacent urea molecules displayed (*green*).

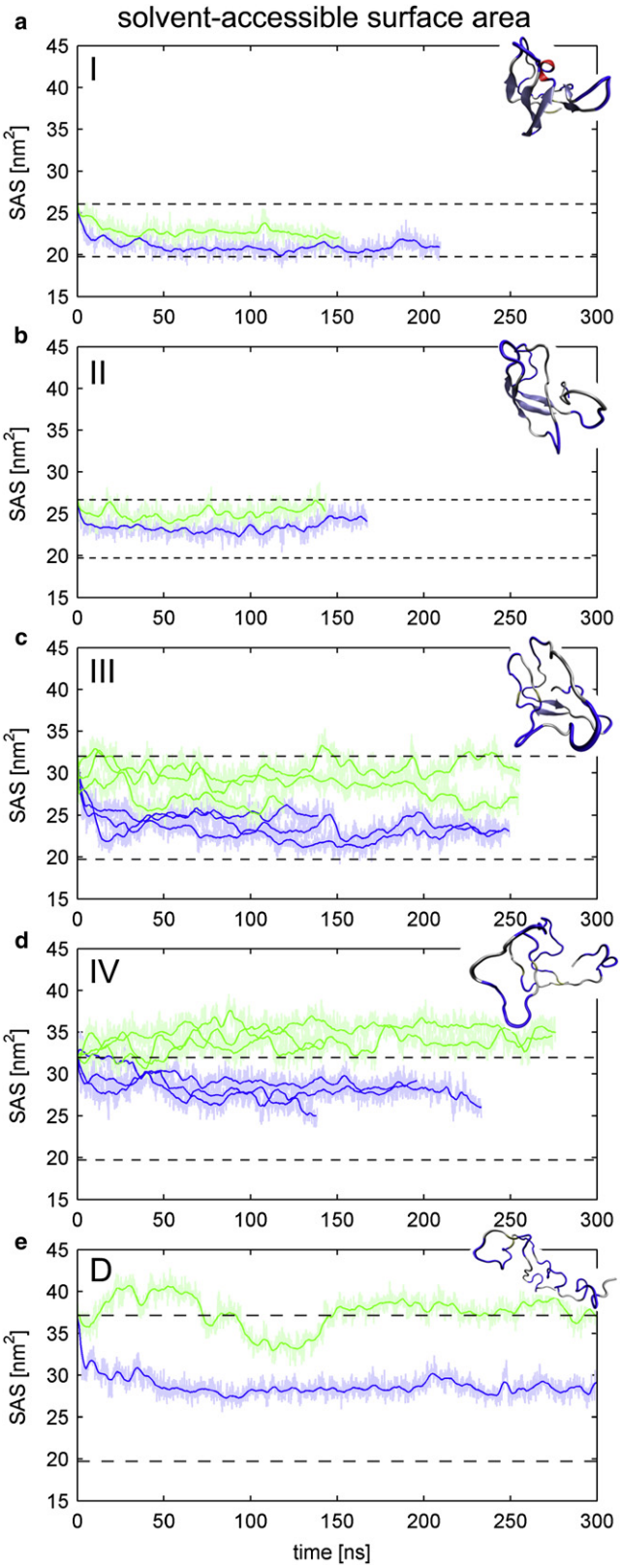

FIGURE 4 SAS of the partially unfolded structures *I–IV*, and *D* (panels *a–e*). The lower dashed lines indicate the SAS of the protein in the crystal structure. The upper-dashed lines show the initial SAS of the respective starting structure.

This observation suggests that urea stabilizes a more-open conformation, which was adopted through thermal fluctuations. This destabilization of less polar contacts by urea might be the first step of unfolding. However, no further destabilizing effects of urea on the protein were observed within the simulation time of ≈450 ns. The contact-destabilization described here is in accordance with results of Lee and van der Vegt (7), who found that urea stabilizes a solvent-separated pair of neopentanes (as a model system for hydrophobic residues), and suggested that urea-separated nonpolar contacts play an essential role in the denaturation process.

Dynamics of partially unfolded structures

To investigate the effect of urea at different states of folding/unfolding, partially unfolded structures (I–IV, D) of the Cold Shock protein were generated in a high-temperature unfolding simulation (Fig. 1), and subsequently simulated in water and in urea at room temperature.

Solvent-accessible surface

The SAS of structures I and II do not show significant differences between water (Fig. 4, *blue lines*) and aqueous urea solution (*green lines*) as solvent. For structure I, the SAS decreases for both solvents from the initial value of ≈26 nm^2 (*upper dashed line*) during the first 40 ns and then fluctuates around constant values which differ from each other by $\Delta_{UW}SAS = 1.7 \pm 0.2\ nm^2$. Structure II shows a similar behavior of the SAS, except that the initial decrease in SAS is less pronounced for both solvents.

Whereas for structures I and II the difference $\Delta_{UW}SAS$ is rather small, a pronounced differential effect of the solvent on the SAS is seen for the more unfolded structures III, IV, and V (Fig. 4, panels *c–e*, respectively). Here, the SAS of the protein remains constant or even increases for most of the simulations with urea, whereas a significant decrease of the SAS is observed for all simulations with water. Correspondingly, as summarized in Table 2, the average $\Delta_{UW}SAS$ is large for the later unfolding stages. In all of these simulations, a hydrophobic collapse is seen for the protein in water, which apparently is prevented by the urea solvent.

Native structure formation

As must be expected, the protein collapse does not lead back to the native state during the few hundred nanoseconds simulation

TABLE 2 Average SAS differences of the protein in the simulations with urea and with water; in all simulations, the SAS of the protein is larger in urea than in water

Starting structure		$\Delta_{UW}SAS$ [nm^2]
N	(0 ps)	0.3 ± 0.1
I	(750 ps)	1.7 ± 0.2
II	(1000 ps)	1.7 ± 0.2
III	(1250 ps)	5.4 ± 0.2
IV	(1500 ps)	5.8 ± 0.2
D	(3000 ps)	8.8 ± 0.3

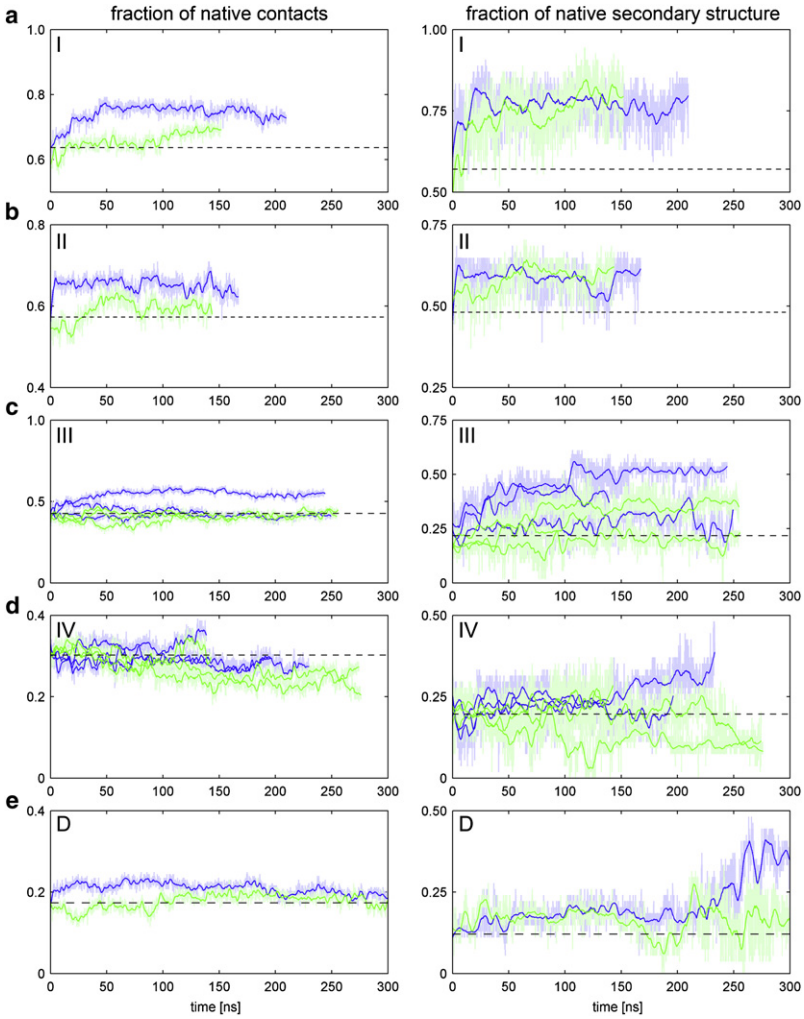

FIGURE 5 Native structure content for the partially unfolded structures *I–IV*, and *D* (*a–e*). (*Left panel*) fraction of native contacts; (*right panel*) fraction of native secondary structure. (*Dashed lines*) Initial values of the respective starting structure.

time, which is still much shorter than measured folding times (32,33). However, for certain structures, partial reformation of native contacts (*left panel* in Fig. 5) and native secondary structure elements (*right panel* in Fig. 5) is observed.

For structures I and II, native contacts form to some extent in water but not in urea (Fig. 5, *a* and *b*), or at least at a slower rate. Hence, the SAS decrease is caused by partial refolding of the protein, which is more pronounced for conformation I (which is closer to the native conformation) than for conformation II. Interestingly, more of the native contacts form in the simulations with water than in those with urea. For structures III, IV, and D (Fig. 5, *c–e*, respectively), no significant differences between the native contact formation for the two different solvents are seen.

Regarding native secondary structure formation, a difference between water and urea is observed for structure III (Fig. 5 *c*), and, less pronounced, for structure IV (Fig. 5 *d*). In these cases, partial native secondary structure is formed in water but not in urea. For the structures I, II, and D (Fig. 5, *a*, *b*, and *e*), no significant difference between both solvents was seen. However, these results are not as clear or statistically certain as those regarding the hydrophobic collapses. Hence, here we focus on the effect of urea on the hydrophobic collapse.

These results are supported by experimental findings. Results from CD-spectroscopy showed that no native state topology is present in collapsed unfolded CsP (34). Moreover, a similar collapse to a disordered state as seen here was observed to precede folding of Cold Shock proteins in FRET experiments (35,36). In particular, as in our simulations, the collapse was found to be faster than secondary structure formation (37). The timescale of the hydrophobic collapses observed here (20–70 ns) is in excellent agreement with the timescales between 50 and 70 ns found in FRET experiments for the Cold Shock protein (38) or other proteins of similar size (37,39).

Collapses at the residue level

Further detailed analyses of the hydrophobic collapse events mentioned above reveal that local collapses of parts of the protein cause the observed SAS decrease. Fig. 6 illustrates an example for this effect. Fig. 6 *a* shows the SAS of Trp^{8} and Ser^{24} of conformation IV during two simulations in water and urea.

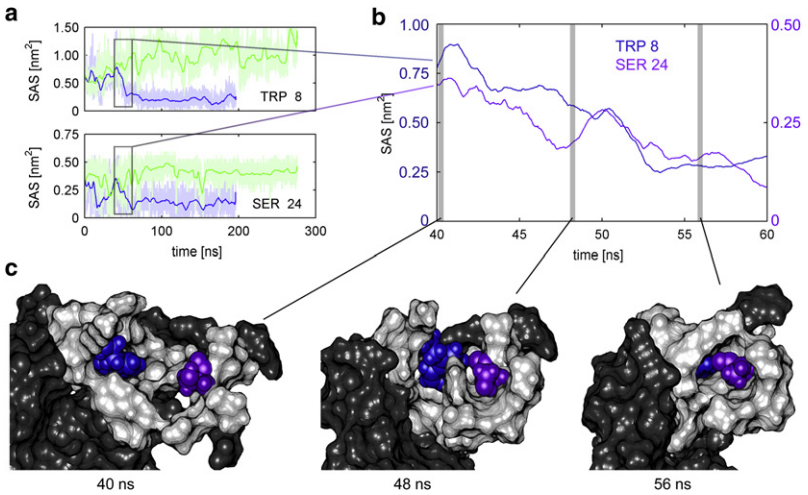

FIGURE 6 Example of a hydrophobic collapse event. (*a*) SAS of Trp^8 and Ser^{24} in simulations of structure *IV* in water (*blue lines*) and urea (*green lines*). (*b*) SAS for both residues in water (*blue*, Trp^8; *magenta*, Ser^{24}). (*c*) Snapshots of the local hydrophobic collapse event with Trp^8 highlighted in blue and Ser^{24} highlighted in magenta. Residues in the collapse region are highlighted in light gray.

As can be seen, in the simulation with water (*blue*), the SAS of both residues decreases, whereas it remains constant or even increases during the simulation with urea (*green*). In Fig. 6 *b*, which shows a zoom of the SAS of both residues in the simulation with water for the time between 40 and 60 ns, a rapid decrease of the SAS to low values is seen between 40 and 55 ns. Fig. 6 *c* shows three snapshots of these two residues and the surrounding protein during this fast local hydrophobic collapse. At the start of the collapse (40 ns, *left panel*), Trp^8 (*blue*) and Ser^{24} (*magenta*) are separated and largely exposed to the solvent, with the adjacent residues (*light gray*) being flexible and loosely packed. After 48 ns, this initially open pocket started closing (*middle panel*), thereby shielding hydrophobic areas from the solvent, and thus reducing the SAS. At 56 ns, the pocket is nearly closed, and Trp^8 and Ser^{24} (as well as neighboring residues) are almost completely buried from the solvent. Similar events are seen for all of the simulations in water and represent the main contribution to the observed total SAS decrease. For urea as solvent, collapses of this kind are rarely seen.

Interactions between urea and the protein

We now address the molecular cause for the impediment of hydrophobic collapses by urea. Previously, favorable apolar contacts of urea with less polar residues were found to be the driving force of denaturation, and it has been suggested that these apolar contacts reduce the hydrophobic effect for less polar residues (8,9,27). Here, a reduced hydrophobic effect in urea solution would indeed nicely explain the absence of hydrophobic collapses. To support this idea, the question needs to be addressed whether the previously observed preferential contacts between urea and less polar residues are also seen in the simulations of the Cold Shock protein. For the folded state, the disruption of an apolar contact between Phe^{27} and the apolar part of Arg^{56} was already shown as one example, and more frequent urea contacts are indeed seen for these and neighboring residues. Now, a more comprehensive and quantitative analysis, in terms of the urea/water contact coefficient C_{UW} of the interactions between urea and all residues of the Cold Shock protein, is presented. Since the focus here is on the interaction of urea with open, noncollapsed, conformations, contact coefficients were calculated from the trajectory of the unfolded structure (D). To assess the influence of protein structure on the contact coefficients, they were also calculated from the trajectory of the folded structure (N).

The contact coefficients for the unfolded state (C^D_{UW}, Fig. 7 *a*) largely agree with those presented previously (8) for glycine-capped tripeptides. In particular, as was expected, less polar residues exhibit strong contact preferences for urea. In addition, the backbone shows strong preference for contacts with urea, whereas urea contacts with charged

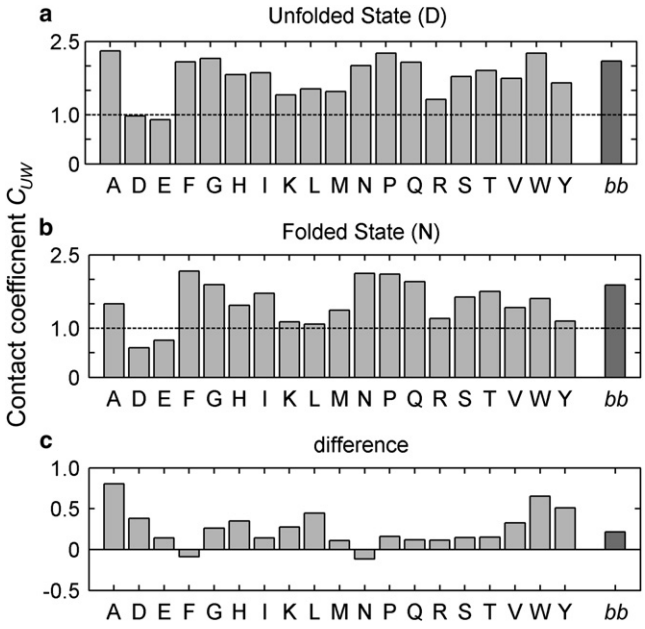

FIGURE 7 Contact coefficient C_{UW} for each residue type and the backbone average (*bb*). (*a*) For the unfolded state (C^D_{UW}), (*b*) for the folded state (C^N_{UW}), and (*c*) difference between unfolded state and folded state (ΔC^{D-N}_{UW}).

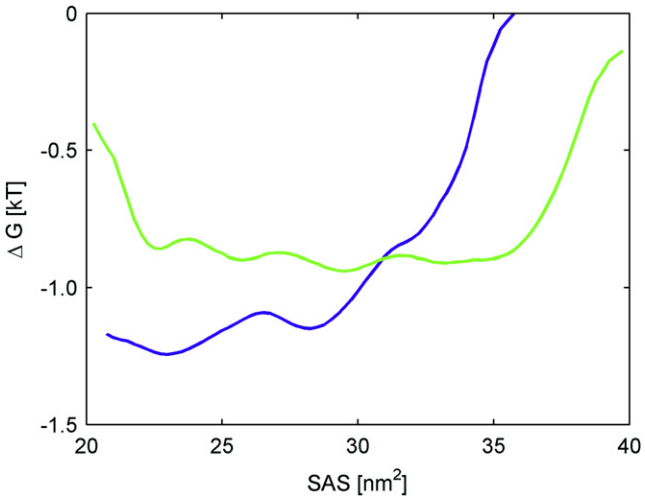

FIGURE 8 The free energy profile of the Cold Shock protein in the two solvents was estimated from a Markov model, with the SAS as reaction coordinate. The estimated free energy of the Cold Shock protein in water (*blue*), and in urea (*green*) is shown.

and polar residues are less frequent. Interestingly, the contact coefficients calculated for the unfolded Cold Shock protein are on average ≈16% higher than those calculated for the tripeptides, which apparently is an effect of the sequence. Other sequence effects on the contact coefficients were also observed. For instance, Gly^{44} has a significantly lower contact coefficient (1.5) than Gly on average (2.1) in the Cold Shock protein, which is due to the presence of the adjacent charged residues Glu^{42}, Glu^{43}, and Glu^{46}. A linear regression of the contact coefficients for the unfolded Cold Shock protein with those for the tripeptides shows a correlation coefficient of $r^2 \approx 0.71$. This suggests that effects from sequence and residual structure in the Cold Shock protein make up ~30% of the residue contact preferences in the unfolded Cold Shock protein.

In contrast, the correlation between the contact coefficients calculated for the folded Cold Shock protein (Fig. 7 *b*) and those of the tripeptides is only $r^2 \approx 0.56$. Further, the correlation between the contact coefficients calculated from the folded and from the unfolded Cold Shock protein is $r^2 \approx 0.74$. These correlations suggest that not only sequence, but also the three-dimensional structure significantly affects the contact coefficients of the individual residues.

Markov model for urea-induced unfolding

Urea was found to impede hydrophobic collapse events of partially unfolded proteins, in particular of more unfolded structures. To further quantify this effect of urea, a Markov model was derived from all simulations of structures N, I–IV, and D, using the SAS as reaction coordinate, which allowed us to estimate the free energy profiles of unfolding/refolding for the two solvents. Fig. 8 shows the free energy profile of the Cold Shock protein in water (*blue line*), and in urea (*green line*). As can be seen, open conformations with a high SAS are energetically unfavorable in water. Hydrophobic collapses reduce the SAS, and, hence, lead to energetically more favorable states. In contrast, no such gradient of the free energy is seen with urea as solvent. Instead, the free energy landscape of the protein in urea is almost flat, in particular even for further unfolded conformations with larger SAS. Therefore, hydrophobic collapses would not lead to a free energy decrease, and hence do not or only rarely occur.

SUMMARY AND CONCLUSION

To investigate how urea drives protein denaturation at the molecular level, we performed extensive molecular dynamics simulations of the Cold Shock protein in both water and aqueous urea solution. To address the question which part of the unfolding process is particularly affected, structures unfolded to different degrees were simulated in both solvents, and the dynamics were compared.

Marked differences were seen already for the native state. Whereas in water, the native state remained unchanged, a larger fraction of the protein surface was exposed to the aqueous urea solution, mainly due to a destabilization of interresidue contacts, e.g., between Phe^{27} and Arg^{56}. Although unfolding of the Cold Shock protein proceeds too slowly (32,33) to occur in our simulations at room temperature, this disruption may well represent primary unfolding steps.

Much more pronounced differences between the two solvents were seen for the partially unfolded conformations. Whereas in water the SAS of the protein decreased markedly toward the value of the native state, much smaller SAS changes were observed in aqueous urea. This effect is more pronounced for simulations starting from later stages of unfolding. For the two most unfolded structures, even an increase of the SAS in urea was seen, indicating further unfolding of the protein. In water, the SAS decrease reflects hydrophobic collapses of the protein as a first refolding step. On the residue level, contacts between residues reform, thus reducing their solvent exposed surface. In urea, no or only slight hydrophobic collapses were observed.

To investigate the molecular origin of this effect, interaction frequencies between urea molecules and the protein residues were analyzed. As expected from previous results for tripeptides (8) and the CI2 protein (9), almost all amino acids were found to show contact preferences for urea. This preference was particularly pronounced for less polar residues as well as for the peptide backbone. This result supports the view that direct interactions of urea with less polar protein parts, rather than polar or indirect interactions, render urea such a good solvent for unfolded peptide chains (16,40). Accordingly, the hydrophobic effect is weaker in aqueous urea solution than in water, which impedes the hydrophobic collapse.

Closer analyses of the differential unfolding kinetics revealed that, unexpectedly, the unfolding rates were hardly affected by urea, whereas the refolding rates were drastically slowed down. This result suggests that urea acts by impeding refolding steps of partially unfolded states and not by actively destabilizing the native state, e.g., by breaking intra-protein hydrogen bonds. The resulting shift in equilibrium was directly seen for the primary unfolding step, and for the later steps inferred from our Markov model. This result agrees with and can explain spectroscopy measurements of the folding and unfolding kinetics of Cold Shock proteins, which show that the (re)folding rate strongly depends on denaturant concentration, whereas the unfolding rate is nearly unaffected (30–32).

Additional analyses of reformation of native contacts and secondary structure elements showed that the dominant effect of urea shifts for the different unfolding stages. Whereas urea impedes the hydrophobic collapse for the more unfolded states, for the more folded structures an inhibition of secondary structure and native contact formation, respectively, is observed. This sequence of prior hydrophobic collapse, followed by secondary structure formation, and finally the formation of native contacts, is in line with previous suggestions about the sequence of folding steps (41).

The emerging picture is that equilibrium fluctuations of the native state, which are reversible in water, are rendered essentially irreversible in urea due to favorable interactions of freshly exposed less polar parts with urea molecules. Accordingly, denaturation proceeds as a ratchetlike sequence of urea-stabilized unfolding fluctuations of the protein.

We thank Ira Tremmel and Ulrike Gerischer for carefully reading the manuscript.

M.C.S. gratefully acknowledges support from the Deutsche Volkswagen Stiftung, grant No. I/78 839.

REFERENCES

1. Muller, N. 1990. A model for the partial reversal of hydrophobic hydration by addition of a urea-like cosolvent. *J. Phys. Chem.* 94:3856–3859.
2. Alonso, D. O. V., and K. A. Dill. 1991. Solvent denaturation and stabilization of globular proteins. *Biochemistry.* 30:5974–5985.
3. Duffy, E. M., P. J. Kowalczyk, and W. L. Jorgensen. 1993. Do denaturants interact with aromatic hydrocarbons in water? *J. Am. Chem. Soc.* 115:9271–9275.
4. Tsai, J., M. Gerstein, and M. Levitt. 1996. Keeping the shape but changing the charges: a simulation study of urea and its iso-steric analogs. *J. Chem. Phys.* 104:9417–9430.
5. Zou, Q., S. M. Habermann-Rottinghaus, and K. P. Murphy. 1998. Urea effects on protein stability: hydrogen bonding and the hydrophobic effect. *Proteins.* 31:107–115.
6. Ikeguchi, M., S. Nakamura, and K. Shimizu. 2001. Molecular dynamics study of hydrophobic effects in aqueous urea solutions. *J. Am. Chem. Soc.* 123:677–682.
7. Lee, M.-E., and N. F. van der Vegt. 2006. Does urea denature hydrophobic interactions? *J. Am. Chem. Soc.* 128:4948–4949.
8. Stumpe, M. C., and H. Grubmüller. 2007. Interaction of urea with amino acids—implications for urea-induced protein denaturation. *J. Am. Chem. Soc.* 129:16126–16131.
9. Stumpe, M. C., and H. Grubmüller. 2008. Polar or apolar—the role of polarity for urea-induced protein denaturation. *PLoS Comput. Biol.* 4:e1000221.
10. Hua, L., R. Zhou, D. Thirumalai, and B. Berne. 2008. Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. *Proc. Natl. Acad. Sci. USA.* 105:16928–16933.
11. England, J. L., V. S. Pande, and G. Haran. 2008. Chemical denaturants inhibit the onset of dewetting. *J. Am. Chem. Soc.* 130:11854–11855.
12. Wallqvist, A., D. G. Covell, and D. Thirumalai. 1998. Hydrophobic interactions in aqueous urea solutions with implications for the mechanism of protein denaturation. *J. Am. Chem. Soc.* 120:427–428.
13. Caballero-Herrera, A., K. Nordstrand, K. D. Berndt, and L. Nilsson. 2005. Effect of urea on peptide conformation in water: molecular dynamics and experimental characterization. *Biophys. J.* 89:842–857.
14. Oostenbrink, C., and W. F. van Gunsteren. 2005. Methane clustering in explicit water: effect of urea on hydrophobic interactions. *Phys. Chem. Chem. Phys.* 7:53–58.
15. O'Brien, E. P., R. I. Dima, B. Brooks, and D. Thirumalai. 2007. Interactions between hydrophobic and ionic solutes in aqueous guanidinium chloride and urea solutions. *J. Am. Chem. Soc.* 129:7346–7353.
16. Tanford, C. 1970. Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv. Protein Chem.* 24:1–95.
17. Berendsen, H. J. C., D. van der Spoel, and R. van Drunen. 1995. GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* 91:43–56.
18. Lindahl, E., B. Hess, and D. van der Spoel. 2001. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* 7:306–317.
19. Jorgensen, W. L., D. S. Maxwell, and J. Tirado-Rives. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 118:11225–11236.
20. Jorgensen, W. L., J. Chandrasekhar, and J. D. Madura. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
21. Smith, L. J., H. J. C. Berendsen, and W. F. van Gunsteren. 2004. Computer simulation of urea-water mixtures: a test of force field parameters for use in biomolecular simulation. *J. Phys. Chem. B.* 108:1065–1071.
22. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an $N\cdot\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.
23. Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, et al. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.
24. Berendsen, H. J. C., J. P. M. Postma, W. F. V. Gunsteren, A. DiNola, and J. R. Haak. 1984. Molecular dynamics with a coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.
25. Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, et al. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235–242.
26. Mueller, U., D. Perl, F. X. Schmid, and U. Heinemann. 2000. Thermal stability and atomic-resolution crystal structure of the *Bacillus caldolyticus* cold shock protein. *J. Mol. Biol.* 297:975–988.
27. Stumpe, M. C., and H. Grubmüller. 2007. Aqueous urea solutions: structure, energetics, and urea aggregation. *J. Phys. Chem. B.* 111:6220–6228.
28. Eisenhaber, F., P. Lijnzaad, P. Argos, C. Sander, and M. Scharf. 1995. The double cubic lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comput. Chem.* 16:273–284.
29. Kabsch, W., and C. Sander. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22:2577–2637.

30. Reid, K. L., H. M. Rodriguez, B. J. Hillier, and L. M. Gregoret. 1998. Stability and folding properties of a model β-sheet protein, *Escherichia coli* CspA. *Protein Sci.* 7:470–479.

31. Schuler, B., W. Kremer, H. R. Kalbitzer, and R. Jaenicke. 2002. Role of entropy in protein thermostability: folding kinetics of a hyperthermophilic cold shock protein at high temperatures using ^{19}F NMR. *Biochemistry.* 41:11670–11680.

32. Schindler, T., M. Herrler, M. A. Marahiel, and F. X. Schmid. 1995. Extremely rapid protein folding in the absence of intermediates. *Nat. Struct. Biol.* 2:663–673.

33. Perl, D., C. Welker, T. Schindler, K. Schroeder, M. A. Marahiel, et al. 1998. Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nat. Struct. Biol.* 5:229–235.

34. Hoffmann, A., A. Kane, D. Nettels, D. E. Hertzog, P. Baumgärtel, et al. 2007. Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy. *Proc. Natl. Acad. Sci. USA.* 104:105–110.

35. Lipman, E. A., B. Schuler, O. Bakajin, and W. A. Eaton. 2003. Single-molecule measurement of protein folding kinetics. *Science.* 301:1233–1235.

36. Magg, C., and F. X. Schmid. 2004. Rapid collapse precedes the fast two-state folding of the Cold Shock protein. *J. Mol. Biol.* 335:1309–1323.

37. Sadqi, M., L. J. Lapidus, and V. Muñoz. 2003. How fast is protein hydrophobic collapse? *Proc. Natl. Acad. Sci. USA.* 100:12117–12122.

38. Nettels, D., I. V. Gopich, A. Hoffmann, and B. Schuler. 2007. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl. Acad. Sci. USA.* 104:2655–2660.

39. Hagen, S. J., and W. A. Eaton. 2000. Two-state expansion and collapse of a polypeptide. *J. Mol. Biol.* 301:1019–1027.

40. Tanford, C. 1968. Protein denaturation. *Adv. Protein Chem.* 23:121–282.

41. Dobson, C. M., P. A. Evans, and S. E. Radford. 1994. Understanding how proteins fold: the lysozyme story so far. *Trends Biochem. Sci.* 19:31–37.