

Dependence of the Size of the Initially Collapsed Form During the Refolding of Barstar on Denaturant Concentration: Evidence for a Continuous Transition

Kalyan K. Sinha and Jayant B. Udgaonkar*

National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bangalore 560065 India

Two-site fluorescence resonance energy transfer (FRET) measurements have been made to determine how two intra-molecular distances contract in the sub-millisecond collapse reaction that occurs initially during the refolding of the small protein barstar. FRET measurements were made on two, single-Cys and single-Trp-containing mutant forms of barstar, Cys25 and Cys62, in each of which a thionitrobenzoate (TNB) adduct was attached to the cysteine thiol. In each protein, the core tryptophan, Trp53, acted as the FRET donor, and the TNB adduct, located either at C25 or at C62, acted as the FRET acceptor. The stabilities as well as observable folding kinetics of the Cys25 and Cys62 mutant proteins were found to be identical. The presence of the TNB adduct on the cysteine did not alter the stability or folding kinetics of either protein. Thus, the FRET-monitored changes in the two labeled mutant proteins, Cys25-TNB and Cys62-TNB, could be compared directly. Refolding was commenced from unfolded protein in 8 M urea, and both the Trp53 to C25-TNB distance and the Trp53 to C62-TNB distance were found to contract upon dilution of urea. The extent of contraction of each distance, which was measured at a few milliseconds of refolding, was dependent continuously on the concentration of urea present during refolding, and was different for the two distances. For either FRET pair, the gradual contraction of distance with a decrease in the concentration of urea in which refolding occurs, was continuous with the contraction of the polypeptide chain that is seen with a decrease in the concentration of urea in the range in which the protein remains completely unfolded. It therefore appears that the products of the initial sub-millisecond refolding reaction of barstar are collapsed forms, whose dimensions do not change cooperatively in an all-or-none manner, but instead, change gradually with a change in concentration of urea. Thus, the sub-millisecond polypeptide chain collapse reaction of barstar upon denaturant dilution, appears to be a continuous structural transition.

© 2005 Elsevier Ltd. All rights reserved.

*Corresponding author

Keywords: collapse; cooperativity; FRET; burst phase; non-cooperative

Introduction

Many proteins appear to fold *via* an initial collapse of the polypeptide chain.^{1–15} The importance of an initial collapse reaction is highlighted in computer simulations of folding,¹⁶ which suggest that folding proceeds through an obligatory, rapidly

collapsed, structure-less globule.^{17–18} Not only does initial collapse of the polypeptide chain quickly reduce the conformational space available for the chain to sample, but it is also expected to facilitate formation of specific structure.¹⁹ Nevertheless, the requirement of a collapse event for correct folding of a polypeptide is poorly understood.^{18,20,21} Proper understanding is further confounded by the observation that for some apparently two-state folding proteins, collapse appears to occur concomitantly with structure formation,^{21–23} while for other such proteins, chain collapse appears to precede structure formation.¹² Clearly, obtaining an

Abbreviations used: TNB, thionitrobenzoate; FRET, fluorescence resonance energy transfer.

E-mail address of the corresponding author: jayant@ncbs.res.in

understanding of the nature and relevance of polypeptide chain collapse during protein folding, remains an important challenge.

A fundamental question concerning the initial polypeptide collapse reaction is whether it is a cooperative two-state transition,^{19,24–25} or whether it is a continuous barrier-less transition occurring through progressively more compact forms as envisaged for homopolymers.^{26–27} The development of new very fast methods for initiating the folding reaction,^{28–30} based on laser-induced temperature-jumps or on microsecond mixing, have indicated that in the case of cytochrome *c*, the initial sub-millisecond collapse of the polypeptide chain from the extended state is a two-state process defined by a free energy barrier.³¹ These results suggested that specific intra-chain interactions develop during initial polypeptide chain collapse and, consequently, that the product of the collapse reaction is a specific intermediate. The observation that the initial collapse reaction during the refolding of apomyoglobin is accompanied by the formation of specific structure supports such a viewpoint,^{32–33} even though for cytochrome *c* itself, very little secondary structure is seen to form during the initial collapse reaction.⁵ On the other hand, other studies indicate that the collapse merely represents a non-specific solvent-dependent transformation, experienced by the unfolded state ensemble upon transfer from a high concentration of denaturant where it is unfolded, to a low concentration of denaturant in which folding is initiated.^{3,34–37}

For several proteins that have not been studied by sub-millisecond kinetic methods, the product of the sub-millisecond collapse reaction manifests itself as a kinetic intermediate at a few milliseconds of folding, where it is relatively amenable to study by many methods, including fluorescence, circular dichroism and amide hydrogen exchange.³⁸ These studies have suggested that these initial intermediates possess some of the specific structure present in the folded protein, at least under strongly stabilizing conditions,³⁹ but have not been able to determine unequivocally whether these intermediates are productive on-pathway intermediates, whether they represent a unique state or an ensemble of forms, and to what extent they possess non-native structure, if any. Moreover, some recent studies have suggested that specific structure might not be present in the initial intermediates of several of the proteins commonly used in folding studies,⁴⁰ and what appear to be intermediates might merely represent the unfolded protein in refolding conditions.^{36–37} Regardless of whether this conclusion is valid, it is clear that a quantitative study of the product(s) of the sub-millisecond polypeptide chain collapse reaction can offer considerable insight into the nature of the collapse reaction itself. Most importantly, characterization of the size of the sub-millisecond folding product(s), and of how the size depends on the concentration of the denaturant in which refolding takes place, will permit the determination of whether the collapse reaction

initiated upon dilution of the denaturant has a cooperative (all-or-none) or continuous dependence on the concentration of urea. Surprisingly, such a study has not been carried out so far for any protein.

The 89 amino acid residue protein barstar, which functions as the intracellular inhibitor of extracellular protease barnase in *Bacillus amyloliquefaciens*, has had its folding pathway studied extensively.^{41–46} The product(s) of the initial (sub-millisecond) refolding reactions manifest themselves as the early intermediate I_E , at a few milliseconds of refolding, which has been shown to be an ensemble of different structural forms.^{43,45} These forms are populated differentially in different environmental conditions,^{39,45} so that the structure of I_E appears different in different conditions of folding. In marginally stable conditions of folding, very little structure is apparent in the product of the initial (sub-millisecond) folding events, when studied at a few milliseconds of refolding, although the product, U_C , is known to be collapsed.¹ It was therefore proposed that the initial product of polypeptide chain collapse is a structure-less globule, U_C , which could represent the unfolded form in refolding conditions. U_C transforms into the structured I_E within a few milliseconds or less, and the extent to which this happens depends on the stability conferred to I_E by the conditions of refolding.³⁹ Moreover, examination of the optical properties of the products of the sub-millisecond collapse reaction by multiple probes, indicated that the $U_C \rightarrow I_E$ reaction occurs in at least two steps, either consecutive or in parallel.³⁹ Thus, it appears that the products of the initial (sub-millisecond)



Figure 1. Structure of barstar. The positions of Trp53 in the core, Ala25 in helix 1 and Leu62 in helix 3 are shown. Trp53 is completely buried in the core, while Ala25 and Leu62 are solvent-exposed. Both Ala25 and Leu62 were mutated independently to Cys, to yield two different single cysteine-containing mutant proteins, Cys25 and Cys62, respectively. The ribbon diagram was generated from PDB file 1btb using Rasmol.⁸⁰

collapse reaction of folding manifest themselves as structure-less U_C and structured I_E , with the relative population of each, as well as the composition of I_E , depending on the stability conferred by the conditions of refolding.

In this study, fluorescence resonance energy transfer (FRET) measurements have been used to determine how two intramolecular distances within the products of the initial (sub-millisecond) collapse reaction of barstar vary as a function of the concentration of urea in which refolding occurs. Two different FRET pairs were studied in two different single-Trp-containing, single-Cys-containing, mutant proteins, both of which contain Trp53 in the core as the FRET donor. One mutant protein (Cys25) has the Cys at position 25, while the other (Cys62) has the Cys at position 62. These residue positions were chosen because both are completely

solvent-exposed in the fully folded protein (Figure 1). Thionitrobenzoate (TNB) was attached to either Cys25 or Cys62, to obtain the labeled proteins Cys25-TNB and Cys62-TNB, and it has been shown that the TNB group at either position quenches the fluorescence of Trp53 by acting as a FRET acceptor.^{47–48} The Trp53 to C25-TNB and Trp53 to C62-TNB distances in the products of the initial (sub-millisecond) collapse reaction of barstar were determined at a few milliseconds of refolding, as a function of the concentration of urea in which refolding took place. It is shown that both distances contract monotonically with a decrease in the concentration of urea, and that this decrease is continuous with the contraction seen with completely unfolded protein upon a decrease in the concentration of urea. The two distances vary differently over a wide range of concentrations of

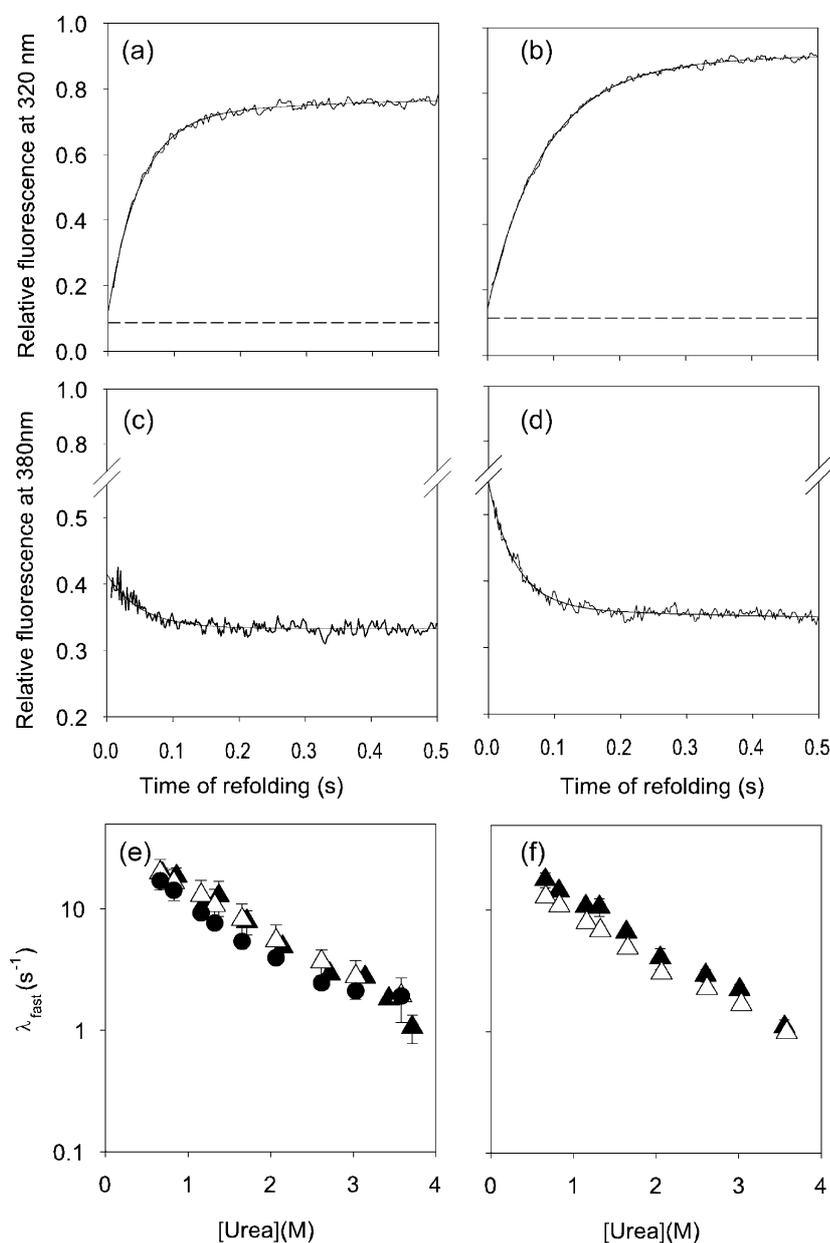


Figure 2. Refolding kinetics of unlabeled and TNB-labeled proteins. The folding reactions of (a) unlabeled Cys25 and (b) unlabeled Cys62, in 0.6 M urea, were monitored by measurement of the increase in Trp53 fluorescence at 320 nm. The folding of (c) Cys25-TNB and (d) Cys62-TNB in 0.6 M urea were monitored by the decrease in fluorescence at 380 nm. Only the first 0.5 s of data are shown. In (a) and (b) the data were normalized to a value of 1 assigned to the value of the fluorescence of the fully folded protein at the end of the folding reactions, and the broken lines denote the relative fluorescence of the unfolded proteins. In (c) and (d), the data were normalized to a value of 1 for protein unfolded in 8 M urea (from which folding was commenced). The continuous lines through the data represent three-exponential fits. The observed fast rate constants of folding of Cys25 and Cys25-TNB are compared in (e), and of Cys62 and Cys62-TNB compared in (f). In (e) and (f), the open triangles represent unlabeled protein whose refolding kinetics were monitored at 320 nm, the filled triangles represent labeled protein whose refolding kinetics were monitored at 380 nm, and the filled circles represent labeled protein whose refolding kinetics were monitored at 320 nm. The error bars represent the standard deviations obtained from three independent repetitions of the experiments.

urea. These results suggest that the sub-millisecond contraction of the polypeptide upon denaturant dilution is not an all-or-none (two-state) process, but is, instead, a gradual process.

Results

Refolding kinetics of unlabeled and TNB-labeled proteins

Figure 2(a) and (b) show representative kinetic traces of the first 0.5 s of refolding in 0.6 M Cys25

and Cys62, respectively. No burst phase increase in Trp53 fluorescence is observed for either protein: the kinetic traces extrapolate to the values of fluorescence corresponding to those of fully unfolded protein in 8 M urea from which the folding reactions were commenced.

Figure 2(c) and (d) illustrate representative kinetic traces of the first 0.5 s of folding in 0.6 M urea of Cys25-TNB and Cys62-TNB, respectively. Folding was monitored by measurement of fluorescence at 380 nm.⁴⁸ For Cys25-TNB, ~55% of the decrease in fluorescence at 380 nm, which accompanies refolding initiated by a jump in

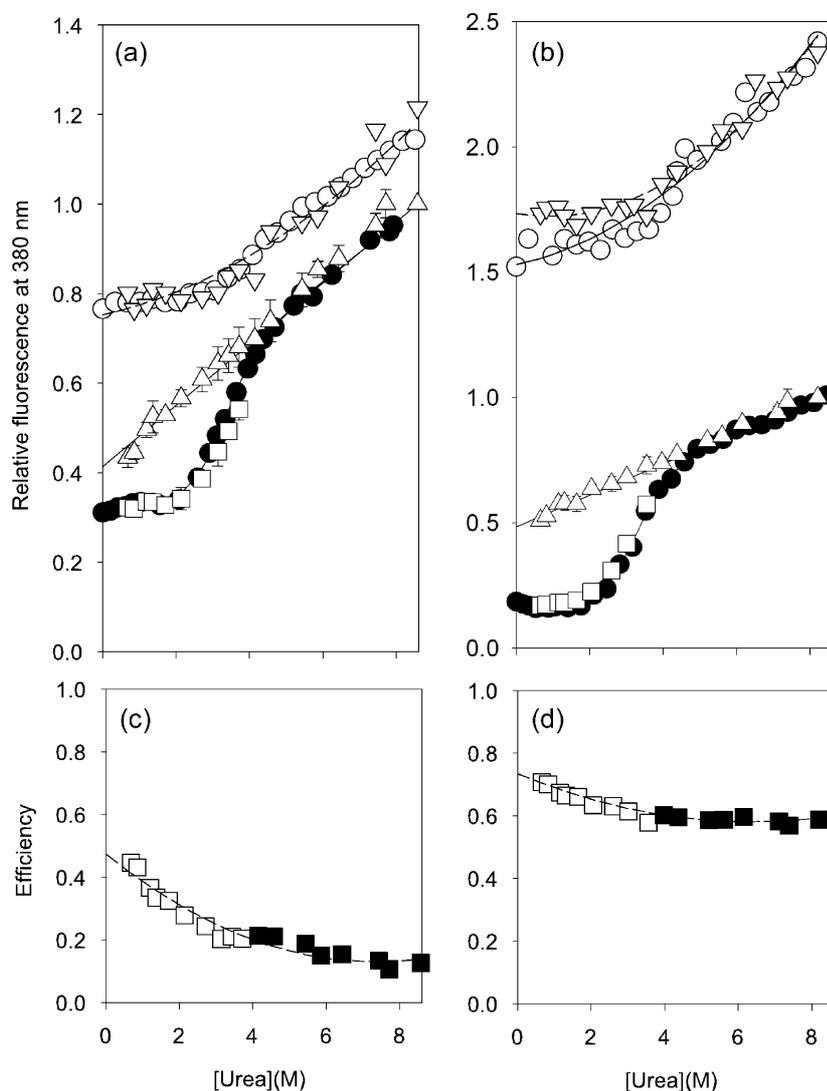


Figure 3. Urea-dependences of the initial phase in folding monitored at 380 nm. Kinetic amplitudes are plotted *versus* equilibrium amplitudes of refolding monitored at 380 nm for (a) Cys25 and Cys25-TNB, and (b) Cys62 and Cys62-TNB. In (a) and (b), open circles represent the equilibrium unfolding transitions of the unlabeled proteins, and inverted triangles represent the signals obtained by extrapolation of the observed kinetic curves of unlabeled protein to $t=0$, and the broken lines represent a fit of the urea-dependence of the fluorescence signal of unlabeled protein to a second-order polynomial. Filled circles represent equilibrium unfolding transitions of the labeled proteins, squares represent the signals corresponding to $t=\infty$ points of kinetic refolding traces of labeled protein, triangles represent the signals obtained by extrapolation of the observed kinetic curves of labeled protein to $t=0$, continuous lines represent fits of the equilibrium unfolding data of labeled protein to a two-state $N \leftrightarrow U$ model for unfolding, and dotted lines are linear extrapolations of the unfolded protein baselines of the labeled protein. In (a) and (b), all data for labeled and unlabeled protein were normalized to a value of 1 for the signal of the labeled protein in 8.3 M urea. The error bars represent standard deviations

obtained from three independent repetitions of the experiments. In (c) and (d) the FRET efficiencies in the forms of each protein, which are populated before commencement of the fast folding phase, are plotted (as open squares) as a function of the concentration of urea for Cys25-TNB and Cys62-TNB, respectively. Also shown (as filled squares) are the FRET efficiencies in the completely unfolded forms of each protein, which are populated in urea concentrations that define the unfolded protein baseline of the equilibrium unfolding transition. For each protein, the FRET efficiency for the form (product of the initial refolding phase) populated before commencement of the fast folding phase at each concentration of urea was determined using equation (1). The value of the fluorescence at 380 nm for unlabeled protein, was taken as F_D . The value of the fluorescence at 380 nm for the corresponding TNB-labeled protein, which was obtained either from extrapolation of the observed fast kinetic phase to $t=0$ (for low concentrations of urea in which the fast phase of folding could be observed), or directly from the equilibrium unfolding transition (for concentrations of urea corresponding to the unfolded protein baseline of the transition), was taken as F_{DA} .

concentration of denaturant from 8 M urea to 0.6 M urea, occurs in the sub-millisecond time domain, too fast to be measured. For Cys62-TNB, ~45% of the increase in FRET efficiency occurs in the unobservable sub-millisecond time domain.

Figure 2(e) and (f) show that the observed fast rates of folding, at any concentration of urea, are the same for Cys25 and Cys25-TNB, as well as for Cys62 and Cys62-TNB. All four unlabeled and labeled proteins fold at essentially the same rate at any concentration of urea. This result was expected

from previous results,⁴⁸ which had shown that all four proteins had identical stabilities, and also unfolded at identical rates.

Initial changes upon denaturant dilution in the fluorescence values at 380 nm of unfolded labeled proteins have linear dependences on the concentration of denaturant

Figure 3(a) compares the kinetic and equilibrium amplitudes of refolding, measured as the changes in

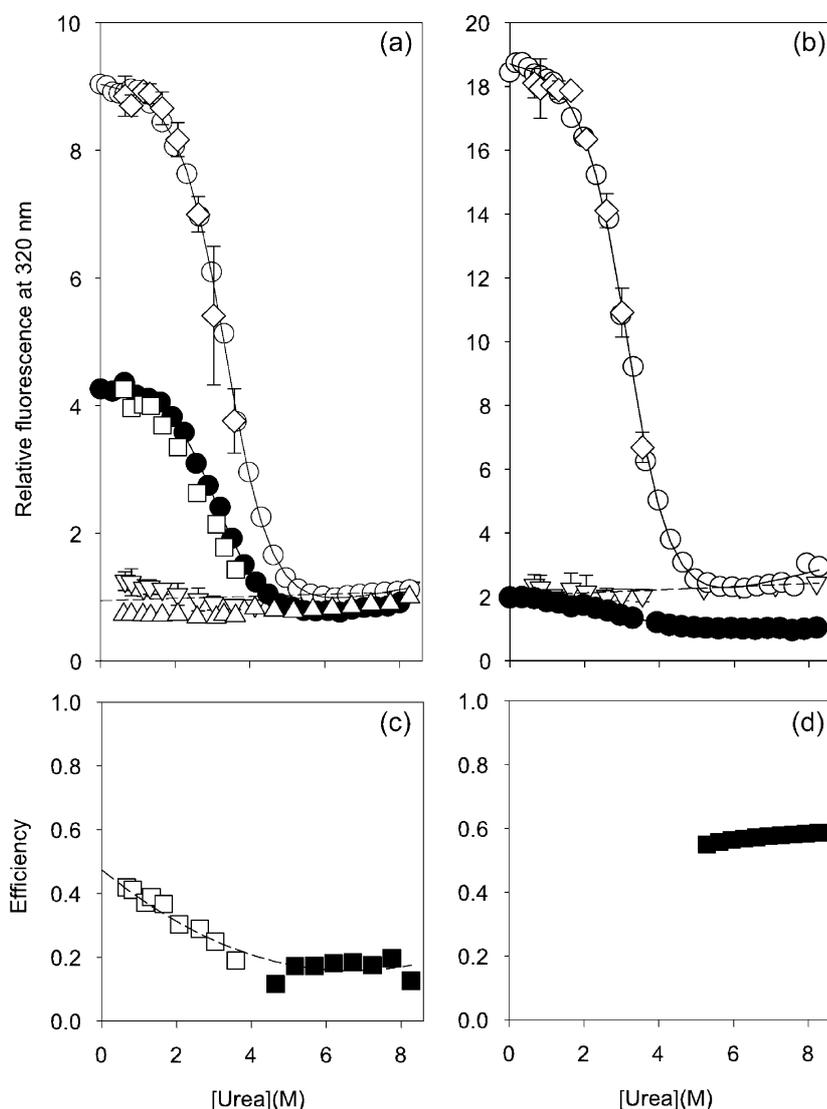


Figure 4. Urea-dependences of the initial phase in folding monitored at 320 nm. Kinetic amplitudes are plotted versus equilibrium amplitudes of refolding monitored at 320 nm for (a) Cys25 and Cys25-TNB, and (b) Cys62 and Cys62-TNB. In (a) and (b), open circles represent the equilibrium unfolding transitions of unlabeled protein, diamonds represent the signals corresponding to $t = \infty$ points of kinetic refolding traces of unlabeled protein, and inverted triangles represent the signals obtained by extrapolation of the observed kinetic curves of unlabeled protein to $t = 0$. Filled circles represent the equilibrium unfolding transition of labeled protein, squares represent the signals corresponding to $t = \infty$ points of kinetic refolding traces of labeled protein, and triangles represent the signals obtained by extrapolation of the observed kinetic curves of labeled protein to $t = 0$. Continuous lines represent fits of the equilibrium unfolding data of unlabeled protein as well as of labeled protein to a two-state $N \leftrightarrow U$ model for unfolding, and broken lines are linear extrapolations of the unfolded protein baselines of the unlabeled protein. In (a) and (b), all data were normalized to a value of 1 for the signal of the labeled protein in 8.3 M urea. The error bars represent standard deviations obtained from three independent repetitions of the experiments. In

(c) and (d) the FRET efficiencies in the forms of each protein, which are populated before commencement of the fast folding phase, are plotted (as open squares) as a function of the concentration of urea for Cys25-TNB. Also shown (as filled squares) are the FRET efficiencies in the completely unfolded forms of each protein, which are populated in urea concentrations that define the unfolded protein baseline of the equilibrium unfolding transition. For each protein, the FRET efficiency for the form (product of the initial refolding phase) populated before commencement of the fast folding phase at each concentration of urea was determined using equation (1). The value of the fluorescence at 320 nm for unlabeled protein, was taken as F_D . The value of the fluorescence at 320 nm for the corresponding TNB-labeled protein, which was obtained either from extrapolation of the observed fast kinetic phase to $t = 0$ (for low concentrations of urea in which the fast phase of folding could be observed), or directly from the equilibrium unfolding transition (for concentrations of urea corresponding to the unfolded protein baseline of the transition), was taken as F_{DA} .

fluorescence at 380 nm, for Cys25 as well as for Cys25-TNB. Figure 3(b) does likewise for Cys62 and Cys62-TNB. For either unlabeled protein, the equilibrium values of fluorescence at 380 nm do not show a co-operative or sigmoidal dependence on the concentration of urea; instead, a smoothly increasing dependence on the concentration of urea is seen, which is similar for both proteins. For either protein, the fluorescence at 380 nm increases by about 30% over the entire range of urea concentration. For Cys25, the $t=0$ points of kinetic refolding curves fall essentially on the equilibrium points, while for Cys62, the $t=0$ points of the kinetic refolding curves show a small deviation at a low concentration of urea.

For either labeled protein, a significant fraction of the initial decrease in fluorescence, which accompanies refolding from 8 M urea, occurs in an unobservable phase, in the sub-millisecond time domain. This fraction is seen to increase linearly with a decrease in the concentration of urea. This linear dependence of the initial sub-millisecond change in fluorescence on the concentration of urea is seen to be exactly the same as that of the linear dependence of the fluorescence of completely unfolded protein on the concentration of urea, and the two linear dependences are seen to be linear extrapolations of each other. It should be noted that the linear dependence of fluorescence on the concentration of urea observed for Cys25-TNB (Figure 3(a)) is different from what is observed for Cys62-TNB (Figure 3(b)).

Dependence of the FRET efficiency on the concentration of urea

The data in Figure 3(a) and (b) were used to calculate the FRET efficiencies of the products of the initial sub-millisecond phase of refolding, as well as of completely unfolded protein, at each concentration of urea. Calculations of the FRET efficiencies by use of equation (2) is straightforward: at each concentration of urea, the fluorescence at $t=0$ of folding of the unlabeled protein was taken as F_D , and the corresponding fluorescence at $t=0$ of folding of the labeled protein was taken as F_{DA} . Figure 3(c) shows how the FRET efficiency of the product of the initial sub-millisecond phase of folding as well as of completely unfolded protein, decreases with an increase in the concentration of urea in the case of Cys25-TNB. Figure 3(d) does likewise for Cys62-TNB. For either protein, it is seen that the FRET efficiencies of the products of the initial sub-millisecond phase of folding, and of completely unfolded protein decrease in a continuous monotonic manner over a wide range of concentrations of urea.

Dependence on the concentration of denaturant of the initial changes in the fluorescence values at 320 nm of unfolded labeled proteins upon denaturant dilution

Figure 4(a) compares the kinetic and equilibrium amplitudes of refolding, measured as the changes in

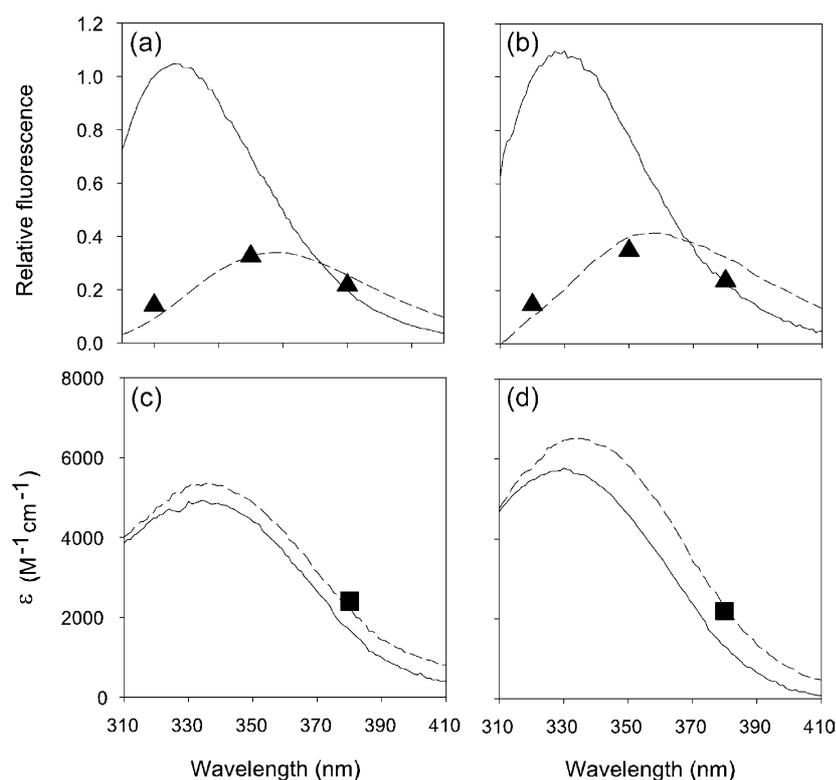


Figure 5. Spectral properties of the product(s) of the initial phase of refolding. The fluorescence spectra of the N state (continuous line) and of the U form in 8 M urea (broken line) are shown together with the values of the fluorescence intensities of the product(s) of the initial phase of refolding at 320 nm, 350 nm and 380 nm (filled triangles), measured at 2 ms of refolding in 0.6 M urea, for (a) Cys25 and (b) Cys62. All fluorescence values were normalized to a value of 1 for the fluorescence intensity of the native protein at 320 nm. The absorbance spectra of the N state (continuous line) and of the U form in 8.3 M urea (broken line) are shown together with the values of the absorption of the product(s) of the initial phase of refolding at 380 nm (filled squares), measured at 10 ms of refolding in 0.6 M urea, for (c) Cys25-TNB and (d) Cys62-TNB.

fluorescence at 320 nm, for Cys25 as well as for Cys25-TNB. Figure 4(b) does likewise for Cys62. For either unlabeled protein, it is observed that the kinetic amplitudes of the refolding experiments account for essentially the complete equilibrium amplitudes: there is no sub-millisecond burst phase change in fluorescence.

When the folding of Cys25-TNB is monitored at 320 nm (Figure 4(a)), again the kinetic amplitudes account for essentially the entire equilibrium amplitude of refolding. The rates of refolding are also identical with the rates measured for Cys25 at 320 nm or for Cys25-TNB at 380 nm (Figure 2(e)). The refolding of Cys62-TNB could not be determined reliably by measurement of fluorescence change at 320 nm, because the total equilibrium change during refolding is itself very small, about 5% of the total equilibrium change observed in the case of Cys62.

The data in Figure 4(a) were used to determine the efficiency of energy transfer in the products of the initial sub-millisecond phase of refolding of Cys25-TNB, as well as of completely unfolded protein, by the use of equation (2). It is observed in the case of Cys25-TNB, that the efficiency decreases continuously with an increase in the concentration of urea, over a wide range of concentrations. It is observed also that the value of efficiency of energy transfer at each concentration of urea, when determined from fluorescence measurements at 320 nm (Figure 4(c)) is similar to the value determined from fluorescence values at 380 nm (Figure 3(c)). In the case of Cys62-TNB, fluorescence measurements at 320 nm allowed efficiencies of energy transfer to be determined only for unfolded protein at high concentrations of urea (Figure 4(d)). These values determined for unfolded Cys62-TNB are in good agreement with values of efficiency determined from fluorescence measurements at 380 nm.

Figures 3(a) and (b), and 4(a) and (b) also bring out the important point, as reported earlier,⁴⁸ that the stabilities of the proteins, labeled as well as unlabeled, are identical. The midpoints of the equilibrium unfolding transitions, the C_m values, are $3.2(\pm 0.1)$ M for Cys25, $3.19(\pm 0.1)$ M for Cys62, $3.2(\pm 0.1)$ M for Cys25-TNB and $3.29(\pm 0.1)$ M for Cys62-TNB.

Spectral characterization of the products of the initial phase of refolding

In Figure 5(a) and (b), it is shown that for both Cys25 and Cys62, the fluorescence emission spectrum of the product of the initial sub-millisecond phase of folding must overlap substantially with that of completely unfolded protein. This observation is in agreement with a previous report, which had shown that the fluorescence emission spectrum of the product of the initial phase product of refolding of *wt* barstar was essentially U-like.⁴⁵ The data in Figure 5(a) and (b) therefore indicate, for Cys25 as well as for Cys62, that (1) the quantum

yield, Q_D , of the donor fluorophore (Trp53), which is proportional to the area under the fluorescence emission spectrum, does not change during the initial (sub-millisecond) phase of refolding, and (2) the emission spectrum of the donor fluorophore does not shift. The observation that the fluorescence emission spectrum of the Trp53 is essentially the same in the initial refolding product and in the U form indicates that Trp53 is as solvated in the product of the initial (sub-millisecond) phase of refolding as it is in the U form. It suggests also that the local motion of the Trp53 side-chain is as unrestricted in the product of the initial refolding phase as it is in the U form.

In Figure 5(c) and (d), it is shown that for both Cys25-TNB and Cys62-TNB, the absorption spectra of the TNB group in folded and unfolded protein overlap to a very large extent. This observation was expected, given that both Cys25 and Cys62 are expected to be as solvent-accessible as are the residues they replace in *wt* barstar. Thus, it is not surprising for either protein, that the wavelengths of maximum absorption of the N and U forms are very similar, indicating that the TNB group is nearly as fully solvated in the N state as it is in the U form. The value of the absorption at 380 nm of the product of the initial sub-millisecond phase of folding is identical with that seen for completely unfolded protein. Hence, the data in Figure 5(c) and (d) indicate the absorption spectrum of the acceptor group (TNB) does not change significantly during the initial sub-millisecond phase of refolding, and indeed, during the entire refolding reaction.

Taken together the data in Figure 5 indicate that, for Cys25-TNB as well as for Cys62-TNB, (1) the overlap integral J (see equation (5)), which is determined as the area that overlaps the emission spectrum of the donor, Trp53, in unlabeled protein, and the absorption spectrum of the acceptor, TNB, in the labeled protein, does not change to any insignificant extent during the initial sub-millisecond phase of refolding, (2) the value of J for the product(s) of the initial sub-millisecond phase of refolding is the same as that for completely unfolded protein, (3) both donor (Trp53) and acceptor (TNB) are as fully solvated in the product of the initial phase of refolding, as they are in the U form, making it likely that both donor and acceptor have unrestricted motional freedom in both forms.

Contraction of donor–acceptor distances in Cys25-TNB and Cys62-TNB during the initial phase of refolding has a continuous dependence on the concentration of urea

To determine the distance separating a donor and an acceptor, using FRET, it is necessary to first determine the value of the Förster distance, R_0 , for each donor–acceptor pair. The value of R_0 can be determined using equation (4), provided the values of the donor quantum yield, Q_D , the spectral overlap integral J , the refractive index, and the orientation factor, κ^2 , are first determined.

Table 1. Parameters used to calculate values of R_0 for the two FRET pairs pairs

Protein	N			U			Burst phase product		
	$J \times 10^{13}$	Q_D^a	R_0 (Å)	$J \times 10^{-13}$	Q_D^a	R_0 (Å)	$J \times 10^{-13}$	Q_D^a	R_0 (Å) ^b
Cys25	5.0	0.27	25.1	5.6	0.11	21.4 ^b	5.6	0.11	21.4 ^b
Cys62	5.7	0.27	25.6	6.5	0.11	21.9	6.5	0.11	21.9

^a Determined from unlabeled protein.^{47–48}

^b Assumed to be same as R_0 of U, because the burst phase product has a U-like fluorescence emission spectrum, and the absorption spectrum of the acceptor TNB group is not significantly different for N and U.

For both the N state and the U form of Cys25-TNB as well as of Cys62-TNB, the values of Q_D had been determined previously,⁴⁸ and are listed in Table 1. The values of J for both the N and U forms of Cys25-TNB as well as of Cys62-TNB, were determined by using equation (5) to determine the spectral overlap between the fluorescence emission spectrum of Trp53 in an unlabeled protein and the absorption spectrum of TNB in the corresponding labeled protein. The use of a value of 2/3 for κ^2 in the case of both the N and U forms, was justified earlier;^{47–48} both donor and acceptor appear to have unrestricted motional freedom, as described above. The values obtained for Q_D and J for the N and U forms of both Cys25-TNB and Cys62-TNB, as well as the values calculated for R_0 , are shown in Table 1.

For the product of the initial sub-millisecond phase of refolding, of Cys25-TNB as well as of Cys62-TNB, the value of R_0 for the U form was used. This use was valid, because (1) the value of Q_D for the product of the initial phase, has been

shown to be the same as that for U (see above), (2) the values of J are similar for the product of the initial (sub-millisecond) refolding phase, and the U form: the emission spectra of Trp53 in both are the same (Figure 4(a) and (b)), and the absorption spectra of TNB in both must be the same, as they are for the N and U forms (Figure 4(c) and (d)), and (3) the values of κ^2 for the burst phase product and the U form must both be equal to 2/3 because both the donor Trp53, as well as the acceptor, TNB, are as fully solvated in the collapsed product of the initial phase of refolding, as they are in the U form (Figure 5, see above). It should be noted also that R_0 has a sixth-root dependence^{49–50} on J and Q_D as well as on κ^2 (equation (4)).

Using the values for R_0 given in Table 1, for the products of the initial phase of refolding of Cys25-TNB and Cys62-TNB, the dependence of FRET efficiency on distance separating donor and acceptor is first determined (Figure 6(a) and (c)) by use of equation (2). Then, using the FRET

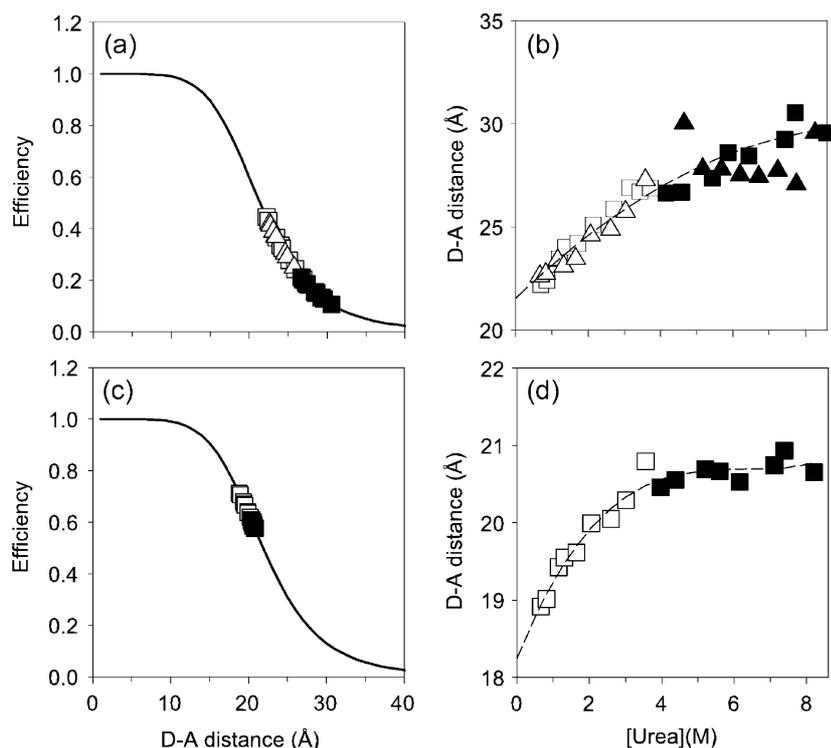


Figure 6. Sub-millisecond collapse of the polypeptide chain upon denaturant dilution. The dependence of FRET efficiency (E) on donor-acceptor separation (R), using a value for R_0 of 21.4 Å in the case of the product of initial refolding of (a) Cys25-TNB, and a value for R_0 of 21.9 Å in the case of product of initial refolding of (c) Cys62-TNB, is plotted according to equation (3) (continuous lines). The FRET efficiencies from Figure 3(c) and (d) are shown as open squares, and the distances separating donor and acceptor can be read off the continuous lines. For each protein, it can be seen that the efficiencies determined from Figure 3(c) or (d) fall on that part of the E versus R curve where E is most sensitive to R (the values determined for R fall in the range $0.5 R_0 < R < 1.5 R_0$). In (b) and (d) the distance separating donor and acceptor in the product(s) of burst phase refolding (open squares), as well as in completely unfolded protein (filled squares), is plotted against concentration of urea.

efficiencies determined in Figure 3(c), (d) and 4(c), the distance of separation corresponding to each value of FRET efficiency, was determined. Figure 6(b) and (d) show how the donor-acceptor separation in the product of the initial phase of refolding varies with the concentration of urea in which refolding is carried out, for Cys25-TNB and Cys62-TNB, respectively. It is seen that for both proteins, the intramolecular distance decreases continuously with a decrease in the concentration of urea, over a broad range of concentrations, and that the fractional change in distance is different for the two intra-molecular distances.

Discussion

Unfolded protein chain dimensions contract continuously with decreasing concentration of denaturant

Very little quantitative information is available about how any unfolded protein chain contracts in volume when the concentration of denaturant used to unfold it is decreased. In this study, FRET measurements have allowed precise measurement of two intramolecular distances within unfolded barstar, which separate the core residue Trp53 from the surface residues C25TNB and C62TNB, and have allowed determination of how these distances within the unfolded protein contract when the concentration of urea is reduced from 8.5 M to 4.5 M. For each distance, the contraction is seen to occur continuously with the decrease in the concentration of urea, indicating that unfolded protein contracts gradually with a decrease in the concentration of urea. Such contraction occurs probably because the polypeptide chain becomes less flexible as intramolecular interactions within the polypeptide chain compete out protein-solvent interactions at lower concentrations of denaturant.⁵¹ Such a contraction in intramolecular distance of the unfolded chain had been apparent as a decrease in the fluorescence lifetime of Trp53 in earlier TR-FRET measurements⁴⁷ of Cys82-TNB distance at high concentrations of denaturant, but the precise extent of contraction was not determined in that study. The increase in density or compaction of unfolded protein with a decrease in the concentration of urea, is likely to be non-specific in the absence of any native-like structure in the unfolded protein.⁵²

It is seen in Figure 6(b) and (d), that the Trp53-Cys25TNB distance appears to saturate at a value of about 30 Å at the highest concentration of urea, and the Trp53-Cys62TNB distance at a value of about 21 Å at the highest concentration of urea. The root-mean-square distance separating two residues in a random coil polypeptide chain separated by n residues has been shown to be equal to $5.8 \times n^{0.49}$ in the absence of an excluded volume effect and $5.7 \times n^{0.58}$ in the presence of an excluded volume effect.⁵³ There is good agreement of the measured

Trp53-Cys62TNB distance with the distance expected (21 Å) when the excluded volume effect is taken into account. On the other hand, the measured Trp53-Cys25TNB distance agrees well with the distance expected (30.5 Å) when the excluded volume effect is not taken into account, and is significantly less than the distance expected (39 Å) when the excluded volume effect is taken into account. It is possible that some local structure in the unfolded protein shortens the Trp53-Cys25TNB distance to below its random coil value, although computer simulations suggest that the presence of such a structure may not affect random coil distances.⁵⁴

Initial sub-millisecond refolding phase of barstar

It is seen in this study as well as in earlier studies^{1,43} that no initial change in intrinsic tryptophan fluorescence is observed at a few milliseconds of refolding of barstar. Nevertheless, it is well known that structural refolding events have occurred at a few milliseconds after transfer from a high concentration to a low concentration of denaturant. The product of the initial submillisecond refolding phase, unlike unfolded protein, is able to bind the hydrophobic dye 8-anilino-1-naphthalene sulfonic acid (ANS).⁴³ It shows the presence of α -helical secondary structure under strongly stabilizing conditions,³⁹ although not under marginally stabilizing conditions.^{1,39} The product of the initial sub-millisecond refolding phase has a higher intrinsic tryptophan fluorescence than that of unfolded protein, under conditions where it is strongly stabilized.^{39,45} Lastly, sub-millisecond events during the refolding of barstar have been measured directly,⁴⁴ and shown to lead to the formation of a loosely compact form. Clearly, the difference in fluorescence intensity between the initially unfolded form and the product of the initial sub-millisecond refolding phase has its origin in the structural changes that occur during the initial refolding phase, which give rise to significant energy transfer between Trp53 and the TNB acceptor. In this study, the dimensions of the product of the initial sub-millisecond refolding phase have been determined over a wide range of concentrations of denaturant.

Contraction of the dimensions of the products of the initial collapse with decreasing concentration of denaturant

The dimensions of the products of the initial sub-millisecond phase of folding are seen to contract in a gradual manner with a decrease in the concentration of urea. For each distance, the contraction observed at a low concentration of urea is continuous with the contraction of unfolded protein, which is observed at a high concentration of urea. This suggests that the collapsed products of the initial phase of refolding, and the unfolded protein respond in an identical manner to a

decrease in the concentration of urea. This is not surprising, because the products of initial refolding, which include U_C and I_E , are known to have their normally buried hydrophobic residues as solvated as they are in the U form, as exemplified by the observation that the fluorescence properties of Trp53 in the products of the initial refolding phase and in the U form are identical (Figure 5).⁴⁵

Since the Trp53-Cys25TNB and Trp53-Cys62TNB distances represent distances of two surface, modified cysteine residues (Cys25-TNB and Cys62-TNB) from a core residue (Trp53), they can be used to estimate fractional change in size with concentration of denaturant. In the case of Cys25-TNB and Cys62-TNB, the distances contract by 17% and 7%, respectively, upon a decrease in concentration of urea from 3 M to 0.6 M. It is not surprising that the distance to Cys62-TNB contracts relatively less than that to Cys25-TNB, given that fewer residues separate Trp53 and C62, than separate Trp53 and C25. It therefore appears that the volume occupied by the polypeptide chain decreases two- to threefold, upon a dilution in urea from 3 M to 0.6 M. Similar relative decreases in volume upon initial collapse have been observed previously for CspTm in single-molecule experiments,⁵⁵ for cytochrome *c* in microsecond small-angle X-ray scattering measurements,⁵⁶ and for barstar in dynamic light-scattering measurements as well as time-resolved measurements of salt-induced non-specific collapse.⁵⁷

Cooperativity of the sub-millisecond folding reactions

The nature of sub-millisecond folding reactions, which appear as burst phase reactions in millisecond measurements, is examined typically by determining how the fluorescence and CD properties of the product of these reactions, when examined at a few milliseconds of refolding, vary upon changing the concentration of denaturant present during refolding. For many proteins, the urea-induced structural transitions of the burst phase product appeared sigmoidal or near-sigmoidal in shape; consequently, the formation of the burst phase products from the U form has been assumed to be a two-state transition. But the evidence that the formation of the burst phase product of a folding reaction may be not two-state, even for an apparently sigmoidal transition, has been mounting. For example, both in the case of thioredoxin⁵⁸ and barstar,³⁹ the dependences of fluorescence and CD properties on the concentration of denaturant in strongly stabilizing conditions, appear to be near-sigmoidal; but they are non-coincident, indicating that formation of the products of sub-millisecond folding reactions is not cooperative. In this context, it should be noted that even the observation of a sigmoidal dependence on the concentration of urea need not necessarily denote an all-or-none transition, but may instead

arise from a gradual transition occurring through a continuum of states.⁵⁹

It is seen in Figures 3(c), (d) and 4(c), that the efficiency in energy transfer varies in a gradual monotonic manner over a wide range of concentrations of urea, and that the dependence of energy-transfer efficiency on the concentration of urea is different for the two distances being monitored. The observed dependences make it very unlikely that the transition between the initially unfolded protein and the product of the initial sub-millisecond refolding phase is a cooperative transition in the case of barstar. In this context, it should be noted that when in previous measurements it was assumed that the initial refolding transition is two-state, then the product of the sub-millisecond refolding phase was found to be only about $0.5 \text{ kcal mol}^{-1}$ more stable than the initially unfolded form.^{39,45} This stabilization is less than the value of $k_B T$, and it would not be unreasonable for the transition between two forms separated by less than $k_B T$ in energy to be a diffusive thermally driven transition. Clearly, a two-state analysis of the initial sub-millisecond transitions is not appropriate. In this study, it has been shown, by measurement of two intramolecular distances, that the dependence of the dimensions of the product of the initial (sub-millisecond) refolding phase on the concentration of urea, is not sigmoidal but is instead, continuous. Clearly, the formation of the product of the initial phase of folding of barstar is very non-cooperative, and appears to be a gradual, continuous transition occurring through near-degenerate states of intermediate densities.⁶⁰⁻⁶¹

It is not known how generally applicable this result is to other proteins. Indirect experimental evidence supporting the notion that the transition between the unfolded form and the sub-millisecond folding products is highly non-cooperative comes from studies of the unfolding of equilibrium models of such burst phase products, in the case of α -lactalbumin⁶²⁻⁶³ as well as barstar.^{57,64} Gradual unfolding has been observed also for the collapsed form of a destabilized mutant variant of the Engrailed homeodomain,⁶⁵ and the transition from the expanded to the initially collapsed form of CspTm does not appear to occur in a two-state manner.¹²

The recent use of sub-millisecond kinetic methods in conjunction with small-angle X-ray scattering or FRET methods, has made it possible to measure directly, the kinetics of chain collapse. The collapse reaction appears to follow exponential kinetics,³¹ and is usually interpreted to represent a two-state transition from the unfolded to collapsed form. A two-state or all-or-none collapse reaction is expected when a free energy barrier separates the unfolded and collapsed forms. But exponential or near-exponential kinetics (with a short lag phase) are also expected in the absence of a free energy barrier when collapse occurs as a diffusive process through a continuum of structural forms, with the time constant of collapse being determined by the

friction experienced by the diffusing polypeptide chain.^{59,66–67} Unfortunately, current theoretical models are not capable of predicting the kinetics of polymer collapse accurately.³⁰ It is therefore not known why such a large range (100 ns to 100 μ s) of time constants for polypeptide chain collapse has been observed by sub-millisecond kinetic measurements of the folding of different proteins.^{3,14,31,33,54,66} The reason could be that collapse is two-state, and is slowed by free energy barriers of different magnitudes; or that collapse is continuous, and is slowed by different magnitudes of hydrodynamic coupling of the polypeptide chain with the solvent. The entropic cost of chain collapse is presumably compensated for by the gain in entropy of water molecules that are released as at least some hydrophobic residues become partially desolvated. The observation that most proteins fold *via* an initial collapse suggests that, with a few exceptions,^{21,23} the hydrophobicity of a protein chain is sufficient that intra-chain interactions dominate over chain-solvent interactions in refolding conditions, to make the overall process of chain collapse energetically favorable.

Thus, even direct measurement of sub-millisecond kinetics cannot easily establish whether they arise from a two-state or a continuous structural transition. In this study, the choice of two intramolecular distances from the core to two different surface positions, as probes for cooperativity of collapse, has made it possible to demonstrate that the density of the burst phase product decreases in a monotonic, apparently continuous, manner, with a decrease in the concentration of urea, in a gradual structural transition (Figure 7).

Similarity between the homopolymer collapse and polypeptide chain collapse

Several theoretical and computational studies have examined the kinetics of chain collapse of homopolymers, to obtain a better understanding of polypeptide collapse. These studies suggest that homopolymers could be expected to collapse in a continuous stage manner,^{19,26,68} or they suggest multiphase kinetics.^{69–72} The collapse of structure-forming heteropolymers (proteins) is obviously more complicated because of the multitude of specific interactions possible,^{19,73} and it is possible that a free energy barrier separates the extended

and compact conformations.⁷⁴ The work reported here shows that a polypeptide chain can behave like a homopolymer during its earliest response to a change in solvent conditions, and it is likely that this happens because very few specific interactions develop during the initial chain collapse reaction.

Chain collapse, secondary structure formation and core formation

Chain collapse can facilitate formation of secondary structure in several ways. When the chain dimensions have become sufficiently compact as a result of chain collapse, steric constraints will favor configurations that are themselves compact, including the elements of secondary structure.^{58,75–76} Also, when the chain dimensions have become sufficiently compact because of collapse, water will be extruded from the core of the collapsed globule; hence, hydrogen bonds between the main chain and water will need to be replaced by those between segments of the main chain leading to formation of secondary structure. Consolidation (rigidification) of the protein core is expected to form only after collapse occurs to a sufficient extent, so that water is extruded from the vicinity of core residues, and is therefore expected to happen only when at least some secondary structure is formed.

Previously, equilibrium studies of the folding of barstar had shown that chain collapse precedes formation of secondary structure, that secondary structure forms along with the extrusion of water from the core, and that the core consolidates only when some secondary structure has formed.^{8,57} Previous kinetic studies of the folding of barstar had shown that essentially no secondary structure is present in the products of the sub-millisecond refolding reaction for folding at concentrations of denaturant that define the start of the equilibrium unfolding transition (1 M GdnHCl, 2.4 M urea),^{1,39} and that the secondary structure content increases only at lower concentrations of urea.³⁹ This study suggests that the secondary structure is probably induced by sufficient compaction of the polypeptide chain at the low concentrations of urea. Other kinetic studies had also shown that core consolidation occurs only when some (although not much) secondary structure has formed.⁷⁷ It is likely that after extrusion of water following collapse,⁷⁸ structural rearrangements leading to

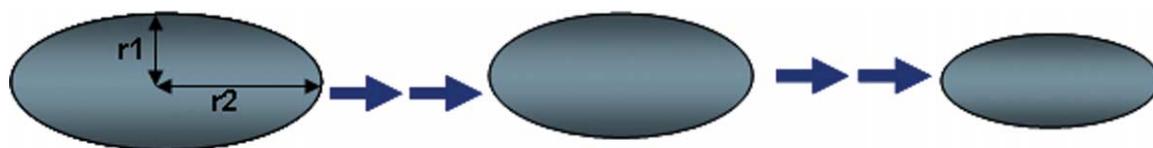


Figure 7. Continuous compaction of the unfolded polypeptide chain occurs upon dilution of urea. Distances r_1 and r_2 represent the C25-TNB to W53 and C62-TNB to W53 distances, respectively. The values of R_1 and R_2 are 29.5 Å and 20.6 Å, respectively, in the unfolded form in 8 M urea, and are 22.2 Å and 18.9 Å, respectively, in the product of the burst phase of refolding in 0.6 M urea. The transition between the expanded and compact forms is shown to be gradual and continuous, and results in approximately a halving of the protein volume.

core consolidation occur in a more cooperative manner because of the necessity that many residue side-chains have to pack tightly. If packing of core residues does indeed occur in an all-or-none manner, while overall collapse occurs gradually, it would explain why probes that report on the core or on changes associated with core formation (such as secondary structure formation), might appear to report cooperative burst phase transitions. It is likely that it is the diffusive, exploratory nature of gradual polypeptide chain collapse that allows the formation of the correct set of intramolecular interactions that enable subsequent consolidation of the core. It is difficult to envisage how an all-or-none (two-state) collapse reaction could achieve this.

Given that the core Trp53 remains fully hydrated in the products of the initial phase of refolding at low ionic strength (Figure 5),⁴⁵ and at all concentrations of urea used for refolding, it is unlikely that core consolidation has occurred in the products of initial collapse, even though some secondary structure forms at the lowest concentrations of urea.³⁹ The absence of a consolidated core, and hence the lack of a defined tertiary structure in the collapsed products of sub-millisecond folding, is probably the reason why the initial collapse transition is a gradual transition. This matter is currently being studied by FRET measurements aimed at measuring distances between buried residues that end up in the core of the fully folded protein.

Materials and Methods

Protein expression, purification and labeling

Wild-type barstar has three tryptophan residues, W38, W44 and W53, and two cysteine residues, C40 and C82. The mutant proteins, W38FW44FC40AC82AA25C (Cys25) and W38FW44FC40AC82AL62C (Cys62), were generated by site-directed mutagenesis, and the proteins purified as described earlier.⁴⁸ The mass of each protein was determined by mass spectrometry on a Micromass Q-TOF Ultima, and was found to be 10,232 Da for Cys25 and 10,190 Da for Cys62, as expected for the proteins when the N-terminal methionine residue remains uncleaved.

The TNB-labeled proteins were obtained by reaction of the protein in 8 M urea at pH 8.5, with a 20-fold molar excess of DTNB (5,5'-dithiobis (2-nitrobenzoic acid)). After the labeling reaction was complete, the free dye and urea were removed by passing the protein through a PD10 column (Pharmacia). Both proteins were found to be >95% labeled, with the expected 196 Da increase in mass due to the presence of the TNB group.

Buffers, solutions and experimental conditions

All reagents used were of the highest purity available from Sigma. The buffer used in all the experiments contained 20 mM Tris and 0.25 mM EDTA at pH 8.0, and all experiments were done at 25 °C. Concentrations of stock urea solutions were determined from measure-

ments of refractive index. In all experiments with unlabeled proteins, the protein concentration used was in the range 5–10 μ M, while in the case of the TNB-labeled proteins, the protein concentration used was \sim 15 μ M.

Determination of protein concentration

For unlabeled proteins, protein concentration was determined by measuring absorbance at 280 nm, using $\epsilon_{280} = 10,000 \text{ M}^{-1} \text{ cm}^{-1}$. Absorbance measurements could not be used directly to determine concentrations of labeled proteins, because of the contribution of the TNB adduct to the absorbance at 280 nm. To determine the extent of this contribution, a solution of labeled protein was divided into two. One part was treated with DTT, while the other part was treated with an equal volume of buffer. Each part was desalted using an AKTA Chromatography system, and the absorbance at 280 nm of each solution was determined. It was found that the presence of the TNB adduct in Cys25-TNB increased its absorbance at 280 nm by 27%, and that the presence of the TNB adduct in Cys62-TNB increased its absorbance at 280 nm by 20%. These effects were taken into account for determination of concentrations of Cys25-TNB and Cys62-TNB solutions by measurements of absorbance at 280 nm.

Equilibrium unfolding experiments

All equilibrium experiments were done using a Fluoromax-3 fluorimeter (Jobin Yvon) or an SFM-4 stopped-flow module (Biologic). The protein samples were incubated in different concentrations of urea for \sim 3 h. Excitation of tryptophan fluorescence was carried out at 295 nm, using a slit-width of 0.4 nm. Emission was monitored at 320 nm and 380 nm for unlabeled protein, and at 380 nm for TNB-labeled proteins, using a slit-width of 10 nm, as described earlier.⁴⁸

Kinetic refolding experiments

All refolding experiments were carried out on either a SFM-4 or SFM-400 stopped-flow module (Biologic, France). Typically, a mixing dead time of 1.8 ms was achieved in experiments with unlabeled proteins by use of the FC-08 cuvette, and a mixing dead-time of 6.2 ms was achieved for TNB-labeled proteins by use of the FC-15 cuvette. Sample excitation was carried out at 295 nm and the emission was monitored at 320 nm for the unlabeled proteins, and at 380 nm for TNB-labeled proteins using +10 nm bandpass filters. A FC-08 cuvette was used for experiments with unlabeled protein, but an FC-15 cuvette was used with labeled proteins because of their significantly lower fluorescence intensities.

To normalize fluorescence values of a labeled protein to that of the corresponding unlabeled protein, the fluorescence values of a known concentration of each protein in 8.3 M urea were determined using the SFM module, using identical excitation and detection parameters.

Acquisition of emission and absorption spectra

All fluorescence emission spectra were acquired on the SFM-400 module coupled with a MOS-250 detection system with both the excitation and emission bandwidth set at 10 nm. The emission spectrum of the product of the burst phase folding reaction was determined by monitoring fluorescence changes at different wavelengths

(320 nm, 350 nm and 380 nm) after initiation of refolding. The $t=0$ points of the kinetic traces obtained after dead-time extrapolation were used to reconstruct a three-wavelength spectrum of burst phase product.

All absorption spectra were acquired on a Cary 100 spectrophotometer using a slit-width of 1 nm, a scan speed of 1 nm/s, and using a cuvette of pathlength 1 cm.

Data analysis

Equilibrium unfolding data

Equilibrium unfolding transitions were analyzed using a two state $N \leftrightarrow U$ model,⁷⁹ to obtain values for the free energy change for unfolding, $\Delta\Delta G_U(H_2O)$, and the midpoint of the unfolding transition (C_m).

Kinetic refolding data

Averages of several kinetic traces were fitted to the sum of three exponentials using the equation:

$$y = y_0 + a_1 e^{-\lambda_1 t} + a_2 e^{-\lambda_2 t} + a_3 e^{-\lambda_3 t} \quad (1)$$

where y is the fluorescence intensity, which changes as a function of time t , and λ_1 , λ_2 and λ_3 are observed rate constants of the different kinetic phases of refolding, and a_1 , a_2 and a_3 are the respective amplitudes.

Analysis of resonance energy transfer

The FRET efficiency for N, U and the initial refolding product was calculated using:

$$E = 1 - \frac{F_{DA}}{F_D} \quad (2)$$

F_{DA} is the fluorescence intensity of the donor in the presence of the acceptor, and is measured for the labeled protein. F_D is the fluorescence intensity of the donor alone in the absence of the acceptor, and is measured for the unlabeled protein. At each concentration of urea, efficiency for the initial refolding product was determined from fluorescence values at $t=0$ fluorescence for the unlabeled (F_D) and labeled (F_{DA}) proteins.

FRET efficiency (E) is related to the distance, R , separating donor and acceptor, by the equation:

$$E = \frac{R_0^6}{R_0^6 + R^6} \quad (3)$$

R_0 is the characteristic Förster's distance for the FRET pair. It was calculated for each FRET (donor-acceptor) pair, in native as well as unfolded protein, using the equation:

$$R_0 = 0.211(Q_D J \kappa^2 n^{-4})^{1/6} \quad (4)$$

Q_D is quantum yield of donor fluorescence, J is the spectral overlap between the emission spectrum of the donor and the absorption spectrum of the acceptor, κ^2 is an orientation factor and n is refractive index of the medium.

Values for Q_D of 0.27 and 0.11 were used for N and U states, respectively, as reported earlier.⁴⁸ J was determined separately for N and U states of each protein using the equation.^{49–50}

$$J = \frac{\int F(\lambda)\epsilon(\lambda)\lambda^4 d\lambda}{\int F(\lambda)d\lambda} \quad (5)$$

To determine J , the fluorescence emission spectra ($F(\lambda)$) of the unlabeled proteins were collected on the SFM 400 coupled to the MOS-250 detection system, using an excitation wavelength of 295 nm (bandwidth=5 nm) and measuring emission from 310 nm to 410 nm (bandwidth=10 nm). Absorption spectra of the TNB-labeled proteins were collected as described above, from 310 nm to 410 nm. All absorption spectra were divided by the respective protein concentration (in M) to obtain $\epsilon(\lambda)$. J was determined as the overlap between $F(\lambda)$ and $\epsilon(\lambda)$ according to equation (5), where the wavelength, λ , is in nm.⁴⁹

A value of 2/3 was used for the value of κ^2 for all calculations,⁴⁷ and values of 1.333 and 1.4 were used for the refractive index of the folded protein in native buffer and unfolded protein in 8 M urea, respectively.

Acknowledgements

We thank G. Krishnamoorthy, M. K. Mathew and B. Rami for discussions. This work was funded by the Tata Institute of Fundamental Research, and by the Department of Science and Technology, Government of India.

References

1. Agashe, V. R., Shastry, M. C. & Udgaonkar, J. B. (1995). Initial hydrophobic collapse in the folding of barstar. *Nature*, **377**, 754–757.
2. Sosnick, T. R., Mayne, L. & Englander, S. W. (1996). Molecular collapse: the rate-limiting step in two state cytochrome c folding. *Proteins: Struct. Funct. Genet.* **24**, 413–426.
3. Chan, C. K., Yi, H., Takahashi, S., Rousseau, D. L., Eaton, W. A. & Hofrichter, J. (1997). Submillisecond protein folding kinetics studied by ultrarapid mixing. *Proc. Natl Acad. Sci. USA*, **94**, 1779–1784.
4. Lorch, M., Mason, J. M., Clarke, J. R. & Parker, J. M. (1999). Effect of core mutations on the folding of a beta-sheet protein: implications for backbone organization in the I-state. *Biochemistry*, **38**, 1377–1385.
5. Akiyama, S., Takahashi, S., Ishimori, K. & Morishima, I. (2000). Stepwise formation of α -helices during cytochrome c folding. *Nature Struct. Biol.* **7**, 514–520.
6. O'Neill, J. C., Jr & Matthews, C. R. (2000). Localized, stereochemically sensitive hydrophobic packing in an early folding intermediate of dihydrofolate reductase from *Escherichia coli*. *J. Mol. Biol.* **295**, 737–744.
7. Arai, M., Ito, K., Inobe, T., Nakao, M., Maki, K., Kamagata, K. *et al.* (2002). Fast compaction of α -lactalbumin during folding studied by stopped flow X-ray scattering. *J. Mol. Biol.* **321**, 121–132.
8. Rami, B. R. & Udgaonkar, J. B. (2002). Mechanism of formation of a productive molten globule form of barstar. *Biochemistry*, **41**, 1710–1716.
9. Buscaglia, M., Schuler, B., Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2003). Kinetics of intramolecular contact formation in a denatured protein. *J. Mol. Biol.* **332**, 9–12.
10. Garcia-Mira, M. M., Sadqui, M., Fischer, N., Sanchoz-Rulz, J. M. & Munoz, V. (2002). Experimental identification of downhill protein folding. *Science*, **298**, 1291–1295.

11. Lipman, E. A., Schuler, B., Bakajin, O. & Eaton, W. A. (2003). Single-molecule measurement of protein folding kinetics. *Science*, **301**, 1233–1235.
12. Magg, C. & Schmid, F. X. (2004). Rapid collapse precedes the fast two-state folding of the cold shock protein. *J. Mol. Biol.* **335**, 1309–1323.
13. Maki, Y., Sasaki, N. & Nakata, M. (2004). Memory effect in the chain-collapse process in a dilute polymer solution. *Macromolecule*, **37**, 5703–5709.
14. Sadqi, M., Lapidus, M. J. & Munoz, V. (2003). How fast is protein hydrophobic collapse? *Proc. Natl Acad. Sci. USA*, **100**, 12117–12122.
15. Enoki, S., Saeki, K., Maki, K. & Kuwajima, K. (2004). Acid denaturation and refolding of green fluorescent protein. *Biochemistry*, **43**, 14238–14248.
16. Dinner, A. R., Sali, A., Sith, L. J., Dobson, C. M. & Karplus, M. (2000). Understanding protein folding via free-energy surfaces from theory and experiment. *Trends Biochem. Sci.* **25**, 331–339.
17. Sali, A., Shakhnovich, E. & Karplus, M. (1994). How does a protein fold? *Nature*, **369**, 248–251.
18. Gutin, A. M., Abkevich, V. I. & Shakhnovich, E. I. (1995). Is burst hydrophobic collapse necessary for protein folding? *Biochemistry*, **34**, 3066–3076.
19. Chan, H. S. & Dill, K. A. (1991). Polymer principles in protein structure and stability. *Annu. Rev. Biophys. Chem.* **20**, 447–490.
20. Noppert, A., Gast, K., Zirwer, D. & Damaschun, G. (1998). Initial hydrophobic collapse is not necessary for folding RNase A. *Fold. Des.* **3**, 213–221.
21. Jacob, J., Krantz, B., Dothager, R. S., Thiyagarajan, P. & Sosnick, T. R. (2004). Early collapse is not an obligate step in protein folding. *J. Mol. Biol.* **338**, 369–382.
22. Schonbrunner, N., Koller, K. P. & Kiefhaber, T. (1997). Folding of disulfide bonded β -sheet protein Tendami-stat: rapid two-state folding without hydrophobic collapse. *J. Mol. Biol.* **256**, 526–538.
23. Plaxco, K. W., Millett, I. S., Segel, D. J., Doniach, S. & Baker, D. (1999). Chain collapse can occur concomitantly with the rate-limiting step in protein folding. *Nature Struct. Biol.* **6**, 554–556.
24. Stigter, D., Alonso, D. O. V. & Dill, K. A. (1991). Protein stability: electrostatics and compact denatured states. *Proc. Natl Acad. Sci. USA*, **88**, 4176–4188.
25. Meisner, K. W. & Sosnick, T. R. (2004). Barrier-limited, microsecond folding of a stable protein measured with hydrogen exchange: implications for downhill folding. *Proc. Natl Acad. Sci. USA*, **101**, 15639–15644.
26. de Gennes, P. G. (1985). Kinetics of collapse for a flexible coil. *J. Phys. Letters*, **46**, L-639–L-642.
27. Eaton, W. A. (1999). Searching for downhill scenarios in protein folding. *Proc. Natl Acad. Sci. USA*, **96**, 5897–5899.
28. Gruebele, M. (1999). The fast protein folding problem. *Annu. Rev. Phys. Chem.* **50**, 485–516.
29. Eaton, W. A., Munoz, V., Hagen, S. J., Jas, G. S., Lapidus, L. J., Henery, E. R. & Hofrichter, J. (2000). Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 327–359.
30. Kubelka, J., Hofrichter, J. & Eaton, W. A. (2004). The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* **14**, 76–88.
31. Shastry, M. C. & Roder, H. (1998). Evidence for barrier-limited protein folding kinetics on the microsecond time scale. *Nature Struct. Biol.* **5**, 385–392.
32. Haruta, N. & Kitagawa, T. (2002). Time-resolved UV resonance Raman investigation of protein folding using a rapid mixer: characterization of kinetic folding intermediates of apomyoglobin. *Biochemistry*, **41**, 6595–6604.
33. Uzawa, T., Akiyama, S., Kimura, T., Takahashi, S., Ishimori, K., Morishima, I. & Fujisawa, T. (2004). Collapse and search dynamics of apomyoglobin folding revealed by sub-millisecond observations of α -helical content and compactness. *Proc. Natl Acad. Sci. USA*, **101**, 1171–1176.
34. Sosnick, T. R., Shtilerman, M. D., Mayne, L. & Englander, S. W. (1997). Ultra fast signals in protein folding and the polypeptide contracted state. *Proc. Natl Acad. Sci. USA*, **94**, 8545–8550.
35. Qi, P. X., Sosnick, T. R. & Englander, S. W. (1998). The burst phase in Ribonuclease A folding and solvent dependence of the unfolded state. *Nature Struct. Biol.* **5**, 882–884.
36. Chen, L., Wildegger, G., Kiefhaber, T., Hodgson, K. O. & Doniach, S. (1998). Kinetics of lysozyme refolding: structural characterization of a nonspecifically collapsed state using time-resolved X-ray scattering. *J. Mol. Biol.* **276**, 225–237.
37. Bhuyan, A. K. & Udgaonkar, J. B. (1999). Relevance of burst phase changes in optical signals of polypeptides during protein folding. In *Perspectives in Structural Biology* (Vijayan, M., Yathindra, N. & Kolaskar, A. S., eds), p. 293, Universities Press, Hyderabad..
38. Juneja, J. & Udgaonkar, J. B. (2003). NMR studies of protein folding. *Curr. Sci.* **84**, 157–172.
39. Pradeep, L. & Udgaonkar, J. B. (2004). Osmolytes induce structure in an early intermediate on the folding pathway of barstar. *J. Biol. Chem.* **279**, 40303–40313.
40. Krantz, B. A., Mayne, L., Rubmbley, J., Englander, S. W. & Sosnick, T. R. (2002). Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. *J. Mol. Biol.* **324**, 359–371.
41. Schreiber, G. & Fersht, A. R. (1993). The refolding of cis- and trans-peptidylprolyl isomers of barstar. *Biochemistry*, **32**, 11195–11203.
42. Shastry, M. C. R., Agashe, V. R. & Udgaonkar, J. B. (1994). Quantitative analysis of the kinetics of denaturation and renaturation of barstar in the folding transition zone. *Protein Sci.* **3**, 1409–1417.
43. Shastry, M. C. R. & Udgaonkar, J. B. (1995). The folding mechanism of barstar: evidence for multiple pathways and multiple intermediates. *J. Mol. Biol.* **247**, 1013–1027.
44. Nolting, B., Golbik, R. & Fersht, A. R. (1995). Submillisecond events in protein folding. *Proc. Natl Acad. Sci. USA*, **92**, 10668–10672.
45. Pradeep, L. & Udgaonkar, J. B. (2002). Differential salt-induced stabilization of structure in the initial folding intermediate ensemble of barstar. *J. Mol. Biol.* **324**, 331–347.
46. Sridevi, K. & Udgaonkar, J. B. (2002). Unfolding rates of barstar determined in native and low denaturant conditions indicate the presence of intermediates. *Biochemistry*, **41**, 1568–1578.
47. Lakshmikanth, G. S., Sridevi, K., Krishnamoorthy, G. & Udgaonkar, J. B. (2001). Structure is lost incrementally during the unfolding of barstar. *Nature Struct. Biol.* **8**, 799–804.
48. Sridevi, K. & Udgaonkar, J. B. (2003). Surface expansion is independent of and occurs faster than core salvation during the unfolding of barstar. *Biochemistry*, **42**, 1551–1563.

49. Lakowicz, J. R. (1999). *Principles of Fluorescence Spectroscopy* Kluwer Academic, Plenum Publishers, New York.
50. Wu, P. & Brand, L. (1994). Resonance energy transfer: methods and applications. *Anal. Biochem.* **218**, 1–13.
51. Krieger, F., Fierz, B., Bieri, O., Drewello, M. & Kiefhaber, T. (2003). Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. *J. Mol. Biol.* **332**, 265–274.
52. Bhavesh, N. S., Juneja, J., Udgaonkar, J. B. & Hosur, R. V. (2004). Native and nonnative conformational preferences in urea unfolded state of barstar. *Protein Sci.* **13**, 3085–3091.
53. Goldenberg, D. P. (2003). Computational simulation of the statistical properties of unfolded proteins. *J. Mol. Biol.* **326**, 1615–1633.
54. Fitzkee, N. C. & Rose, G. D. (2004). Reassessing random coil statistics in unfolded proteins. *Proc. Natl Acad. Sci. USA*, **101**, 12497–12502.
55. Lipman, E. A., Schuler, B., Bakajin, O. & Eaton, W. A. (2003). Single-molecule measurement of protein folding kinetics. *Science*, **301**, 1233–1235.
56. Akiyama, S., Takahashi, S., Kimura, T., Ishimori, K., Morishima, I., Nishikawa, Y. & Fujisawa, T. (2002). Conformational landscape of cytochrome c folding studied by microsecond-resolved small-angle X-ray scattering. *Proc. Natl Acad. Sci. USA*, **99**, 1329–1334.
57. Rami, B. R., Krishnamoorthy, G. & Udgaonkar, J. B. (2003). Dynamics of the core tryptophan during the formation of a productive molten globule intermediate of barstar. *Biochemistry*, **42**, 7986–8000.
58. Georgescu, R. E., Lee, J. H., Goldberg, M. E., Tasayko, M. L. & Chaffotte, A. F. (1998). Proline isomerization-independent accumulation of early intermediate and heterogeneity of the folding pathways of a mixed alpha/beta protein *Escherichia coli* thioredoxin. *Biochemistry*, **37**, 10286–10297.
59. Parker, M. J. & Marqusee, S. (1999). The cooperativity of burst phase reactions explored. *J. Mol. Biol.* **293**, 1195–1210.
60. Dill, K. A. & Shortle, D. (1991). Denatured states of proteins. *Annu. Rev. Biochem.* **60**, 795–825.
61. Finkelstein, A. V. & Ptitsyn, O. B. (2002). *Protein Physics*, pp. 207–237, Academic Press, London.
62. Shimizu, A., Ikeguchi, M. & Sugai, S. (1993). Unfolding of the molten globule state of alpha-lactalbumin studied by ¹H NMR. *Biochemistry*, **32**, 13198–13203.
63. Schulman, B. A., Kim, P. S., Dobson, C. M. & Redfield, C. (1997). A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nature Struct. Biol.* **4**, 630–634.
64. Rami, B. R. & Udgaonkar, J. B. (2002). Mechanism of formation of a productive molten globule form of barstar. *Biochemistry*, **41**, 1710–1716.
65. Mayor, U., Grossmann, J. G., Foster, N. W., Freund, S. M. V. & Fersht, A. R. (2003). The denatured state of Engrailed homeodomain under denaturing and native conditions. *J. Mol. Biol.* **333**, 977–991.
66. Hagen, S. J. & Eaton, W. A. (2002). Two-state expansion and collapse of a polypeptide. *J. Mol. Biol.* **297**, 781–789.
67. Hagen, S. J. (2003). Experimental decay kinetics in “downhill” protein folding. *Proteins: Struct. Funct. Genet.* **50**, 1–4.
68. Williams, C., Brochard, F. & Frisch, H. L. (1981). Polymer collapse. *Annu. Rev. Phys. Chem.* **32**, 433–451.
69. Thirumalai, D. (1995). From minimal models to real proteins: time scales for protein folding kinetics. *J. Phys.* **5**, 1457–1467.
70. Ma, J., Straub, J. E. & Shakhnovich, E. I. (1995). Simulation study of the collapse of linear and ring homopolymers. *J. Chem. Phys.* **103**, 2615–2624.
71. Pitard, E. & Orland, H. (1998). Dynamics of the swelling or collapse of a homopolymer. *Europhys. Letters*, **41**, 467–472.
72. Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry*, **29**, 7133–7155.
73. Kuznetsov, Y. A., Timoshenko, E. G. & Dawson, K. A. (1995). Kinetics of the collapse transition of homopolymers and random copolymers. *J. Chem. Phys.* **103**, 4807–4818.
74. Pande, V. S. J. & Rokhsar, D. S. (1998). Is the molten globule a third phase of proteins? *Proc. Natl Acad. Sci. USA*, **95**, 1490–1494.
75. Chan, H. S. & Dill, K. A. (1990). Origin of structure in globular proteins. *Proc. Natl Acad. Sci. USA*, **87**, 6388–6392.
76. Dill, K. A., Fiebig, K. M. & Chan, H. S. (1993). Cooperativity in protein-folding kinetics. *Proc. Natl Acad. Sci. USA*, **90**, 1942–1946.
77. Sridevi, K., Juneja, J., Bhuyan, A. K., Krishnamoorthy, G. & Udgaonkar, J. B. (2000). The slow folding reaction of barstar: the core tryptophan region attains tight packing before substantial secondary and tertiary structure formation and final compaction of the polypeptide chain. *J. Mol. Biol.* **302**, 479–495.
78. Frieden, C. (2003). The kinetics of side chain stabilization during protein folding. *Biochemistry*, **42**, 12439–12446.
79. Agashe, V. R. & Udgaonkar, J. B. (1995). Thermodynamics of denaturation of barstar: evidence for cold denaturation and evaluation of the interaction with guanidine hydrochloride. *Biochemistry*, **34**, 3286–3299.
80. Lubienski, M. J., Bycroft, M., Freund, S. M. V. & Fersht, A. R. (1994). Three-dimensional solution structure and ¹³C assignments of barstar using nuclear magnetic resonance spectroscopy. *Biochemistry*, **33**, 8866–8877.

Edited by C. R. Matthews

(Received 26 March 2005; received in revised form 15 August 2005; accepted 23 August 2005)
Available online 12 September 2005