

Increasing Stability Reduces Conformational Heterogeneity in a Protein Folding Intermediate Ensemble

K. Sridevi¹, G. S. Lakshmikanth², G. Krishnamoorthy^{2*} and Jayant B. Udgaonkar^{1*}

¹National Centre for Biological Sciences, Tata Institute of Fundamental Research Bangalore 560065, India

²Department of Chemical Sciences, Tata Institute of Fundamental Research Mumbai 400005, India

A multi-site, time-resolved fluorescence resonance energy transfer methodology has been used to study structural heterogeneity in a late folding intermediate ensemble, I_L , of the small protein barstar. Four different intra-molecular distances have been measured within the structural components of I_L . The I_L ensemble is shown to consist of different sub-populations of molecules, in each of which one or more of the four distances are native-like and the remaining distances are unfolded-like. In very stable conditions that favor formation of I_L , all four distances are native-like in most molecules. In less stable conditions, one or more distances are unfolded-like. As stability is decreased, the proportion of molecules with unfolded-like distances increases. Thus, the results show that protein folding intermediates are ensembles of different structural forms, and they demonstrate that conformational entropy increases as structures become less stable. These observations provide direct experimental evidence in support of a basic tenet of energy landscape theory for protein folding, that available conformational space, as represented by structural heterogeneity in I_L , becomes restricted as the stability is increased. The results also vindicate an important prediction of energy landscape theory, that different folding pathways may become dominant under different folding conditions. In more stable folding conditions, uniformly native-like compactness is achieved during folding to I_L , whereas in less stable conditions, uniformly native-like compactness is achieved only later during the folding of I_L to N.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: multi-site FRET; time-resolved fluorescence; protein folding; barstar; conformational heterogeneity

*Corresponding authors

Introduction

The process of protein folding involves the formation of a unique, well-defined native state, from a very large ensemble of unfolded forms. Folding pathways are predicted to resemble a many-dimensional energy landscape with an energy gradient favoring the native state.¹ On the free energy landscape that a protein molecule can explore during folding, numerous conformations

having different energies are sampled, with the most stable one being the final native state.² The funnel-shaped energy landscapes predict that the conformational space becomes narrow as the stability of the structures increases. It has been difficult to demonstrate experimentally the presence of such an energy landscape for protein folding. In particular, establishing the ensemble nature of observed protein folding intermediates, and determining how the structural composition of the observed intermediates changes as the energy bias towards the native state is increased, have been challenges of long standing. Progress in this regard has been slow because of the inherent difficulty in structurally characterizing folding intermediates, and because the many spectroscopic probes typically used to monitor folding reactions,

Abbreviations used: N, native; U, unfolded; FRET, fluorescence resonance energy transfer; TRFRET, time-resolved FRET; MEM, maximum entropy method; TNB, thionitrobenzoate.

E-mail addresses of the corresponding authors: gk@tifr.res.in; jayant@ncbs.res.in

such as circular dichroism and fluorescence intensity, report on ensemble-averaged properties of all molecules present.

To determine structural heterogeneity in an intermediate ensemble, it is necessary to identify the presence of multiple conformations, and to distinguish structurally between them. Fluorescence resonance energy transfer (FRET) methods allow different conformations to be distinguished on the basis of how they differ in a particular intramolecular distance, and measurement of FRET efficiency during folding permits the observation of conformational changes occurring in a specific region of the protein as it folds. Changes in the FRET efficiency between a single donor–acceptor fluorophore pair have been used to monitor the initial steps during the refolding of cytochrome *c*^{3,4} and acyl-CoA binding protein.⁵ More detailed structural analysis is possible with multiple FRET pairs, with each FRET pair engineered to probe the structural change that occurs in a different region of the protein. For example, such a multi-site FRET approach has been used to measure unfolding rates in different regions of barstar as it unfolds.⁶ A change in FRET efficiency, measured by steady-state fluorescence intensity, can, however, give only ensemble-averaged information about the overall increase or decrease in donor–acceptor separation, and cannot give this information for each distinct member of the protein ensemble. This limitation can be overcome by determining FRET using time-resolved fluorescence measurements, in which FRET efficiency is estimated from the fluorescence lifetime of the donor fluorophore in the presence and in the absence of an acceptor fluorophore. In the presence of the acceptor fluorophore, the donor fluorescence lifetime depends on the donor–acceptor distances: long lifetimes correspond to large distances, and short lifetimes to small distances, provided that the relative orientation of the fluorophores does not change.⁷ The fractional amplitude of each lifetime corresponds to the fraction of molecules populating a particular conformation. Thus, time-resolved fluorescence resonance energy transfer (TRFRET) methods^{8–11} allow a distance-based, and hence structure-based, distinction to be made between different conformations when all are present together, and allow each state to be quantified. In this respect, the TRFRET method has a major advantage over other methods of determining size, such as small-angle X-ray scattering,^{12–14} dynamic light-scattering,^{15,16} and time-resolved fluorescence anisotropy decay,^{17–20} all of which rely on a model-based analysis to determine the size of the molecules, and all of which, with the exception of dynamic light-scattering, yield mainly an ensemble-averaged size of all the forms of the protein present.

The power of the TRFRET method is greatly enhanced when the fluorescence intensity decays are analyzed by the model-independent maximum entropy method (MEM).^{7,21,22} The fluorescence life-

time distributions so obtained enable the distinction of sub-populations in an ensemble, and provide an estimate of the conformational heterogeneity in each sub-population. The first reported use of TRFRET coupled to MEM analysis was to characterize the equilibrium unfolding of barstar,⁷ using a single donor–acceptor pair. The TRFRET method showed, for the first time for any protein, that structure was lost in a continuously incremental manner during denaturant-induced unfolding. A similar result was obtained later for the equilibrium unfolding of cytochrome *c*.^{3,23} The only alternative to TRFRET studies coupled to MEM analysis, for identifying heterogeneity in an equilibrium unfolding reaction, are single-molecule FRET studies,^{24–26} but the former are preferable⁷ because they yield data that are statistically more significant. This is the first report describing the use of multi-site TRFRET to characterize the structural composition of a kinetic folding intermediate ensemble.

The 89 residue protein, barstar, which functions as the natural inhibitor to the barnase in *Bacillus amyloliquifaciens*, has been used extensively as a model for protein folding studies. Equilibrium-unfolded barstar comprises of slow-refolding U_S molecules and fast-refolding U_F molecules, populated to 66(\pm 4)% and 34(\pm 4)% at 10 °C, respectively.²⁰ In U_S , the Tyr47–Pro48 peptide bond is in the *trans* configuration, while it is in the *cis* configuration in both U_F and N .^{27,28} The ratio of U_F to U_S in equilibrium-unfolded protein has been shown to be 66:34 for wild-type barstar²⁸ and for several mutant forms that retain Pro48.^{20,27} The major folding reaction is $U_S \rightleftharpoons I_E \rightleftharpoons I_L \rightleftharpoons N$. I_E is an early collapsed intermediate that forms within a few milliseconds, and that folds to the late intermediate I_L within 1 s. U_F molecules too collapse rapidly during folding, and the collapsed form, I_F folds to N within 1 s.

At 10 °C, the overall rate of folding of I_L to N is so slow (\sim 0.06 min⁻¹),¹⁹ that at one minute of refolding, only \sim 6% of I_L has had a chance to fold to N . Nevertheless, double-jump experiments^{20,29} show that \sim 40% of the molecules present at one minute of refolding, unfold at the same rate as does N , suggesting that these 40% of molecules must be N . Indeed, this is expected, because the equilibrium ratio of U_F to U_S is 34:66, and all 34% of the molecules^{20,27,28} originally present as U_F would have completed folding to N within one minute. The remaining molecules are expected to be present either as I_L or as U .

The structural properties of I_L appear to be different under different folding conditions. In 2 M urea and at 10 °C, I_L appears to be compact with a flexible core, and possesses relatively little specific structure.²⁰ In low concentrations of GdnHCl at 25 °C, I_L appears to be far more native-like (it had been called I_N) in its level of secondary and tertiary structure.^{30,31} In low concentrations of urea at 25 °C, I_L is sufficiently native-like to be capable of binding to and inhibiting barnase.²⁷ The

smaller extent of structure in I_L seen in 2 M urea, 10 °C, as compared to that seen in low concentrations of denaturant at 25 °C, correlates well with its stability relative to that of N: at 10 °C, I_L unfolds nearly 25-fold faster than N,³² while at 25 °C, it unfolds only sevenfold faster than N.^{28,29} This is not surprising, because N itself is considerably less stable at 10 °C than at 25 °C,³³ and a compact intermediate such as I_L might be expected to be even more destabilized at the lower temperature. The stability of I_L decreases with increasing concentration of denaturant,^{27–29,32} but denaturant-induced unfolding of I_L is not cooperative: different structural probes yield non-coincident unfolding transitions.³¹ These results as well as the results of real-time NMR experiments³¹ suggested that I_L might be structurally heterogeneous. Nevertheless, different structural components of the I_L ensemble could not be differentiated even with the use of multiple structural probes to monitor the slow $I_L \rightarrow N$ reaction.²⁰

In the present study, multi-site TRFRET measurements have been used to characterize the structural heterogeneity in I_L in terms of four intra-molecular distances, and to characterize how this heterogeneity varies as the folding conditions are made more or less stabilizing. The core tryptophan residue (Trp53) of barstar served as the donor fluorophore, and a Cys residue labeled with a thionitrobenzoate (TNB) moiety served as the acceptor. An acceptor site was engineered at each of four different locations by cysteine substitution mutagenesis, followed by labeling with TNB. The Cys residues were engineered, in four different single-Cys single-Trp mutant proteins, into four different regions: helix 1 (Cys25), helix 2 (Cys40), helix 3 (Cys62) and the third β strand (Cys82). Analysis of the fluorescence lifetimes in the TNB-labeled proteins (Cys25TNB, Cys40TNB, Cys62TNB and Cys82TNB) was used to determine the structural characteristics of I_L and its slow folding to N at 10 °C in different concentrations of urea. Care has to be taken in the analysis of FRET data, and the validity of the analysis of FRET data for the four labeled proteins used in this study was established previously.^{6,7} It was possible to compare directly the folding kinetics of the four different mutant proteins because they have similar stabilities, which are unaltered upon labeling by TNB,⁶ and because the unfolding⁶ and refolding kinetics³² of the unlabeled proteins are similar. The multi-site TRFRET approach has allowed easy identification of conformational heterogeneity in I_L , and has shown how different structural components of the I_L ensemble are populated in conditions that confer different stability.

Results

Different FRET pairs show different amplitude of the slow refolding phase

The refolding reaction of barstar can be moni-

tored by the changes in the FRET efficiency due to changes in the donor–acceptor distances during folding. The fluorescence intensity of the donor reduces upon refolding, due to increased energy transfer to the acceptor, which is closer to the donor in the compact native state. In all the mutant forms of barstar studied here, Trp53 was the donor. Upon excitation at 295 nm, the changes in the fluorescence emission of Trp53 at 380 nm have been shown to represent directly the changes in FRET efficiency during folding.⁶ Figure 1 illustrates the refolding kinetics of Cys25TNB, Cys40TNB, Cys62TNB and Cys82TNB, in 1.4 M urea at 10 °C. Figure 1 also shows (in insets) the fluorescence spectra of the N state, the U form and the species populated at one minute of refolding, for all four labeled proteins. Only in the case of Cys82TNB, is the spectrum collected at one minute of refolding more unfolded-like rather than native-like. The spectra are consistent with the kinetic curves.

The observable changes in fluorescence intensity at 380 nm (FRET efficiency) occur at an apparent rate of $0.06(\pm 0.01) \text{ min}^{-1}$ for all the proteins. At one minute of refolding in 1.4 M urea, double-jump experiments in the case of both Cys40 and Cys82³² show that $\sim 60\%$ of the protein molecules unfold 25-fold faster than do molecules of N, thereby identifying them as I_L . The slow kinetic phase therefore originates from the folding of the 60% molecules present as I_L at its start. Thus, the observed changes in FRET efficiency suggest that during the folding of I_L to N in 1.4 M urea, only small changes occur in the Trp53–Cys25TNB, Trp53–Cys40TNB and Trp53–Cys62TNB distances, while a large change occurs in the Trp53–Cys82TNB distance. In other words, in 1.4 M urea, I_L already appears to be very native-like in terms of the Trp53–Cys25TNB, Trp53–Cys40TNB and Trp53–Cys62TNB distances, but unfolded-like with respect to the Trp53–Cys82TNB distance. The observation that the rate of the $I_L \rightarrow N$ reaction is the same for all four proteins (Figure 1), together with the observation that N is equally stable for all four labeled proteins,⁶ suggests that the overall stability of I_L is the same for all four labeled proteins.

Fractional changes in FRET efficiency, observed by changes in fluorescence intensity, can imply that (1) a fraction of molecules has N-like distances and a fraction has U-like distances, or (2) all molecules have acquired a donor–acceptor distance between that of the N and U states, or (3) an ensemble of molecules having a wide range of donor–acceptor distances, from the compact N-like to the extended U-like, is present. These possibilities were resolved by TRFRET measurements.

Fluorescence lifetimes in the folded and unfolded TNB-labeled proteins

The fluorescence lifetimes obtained by the discrete analysis of the fluorescence intensity decays

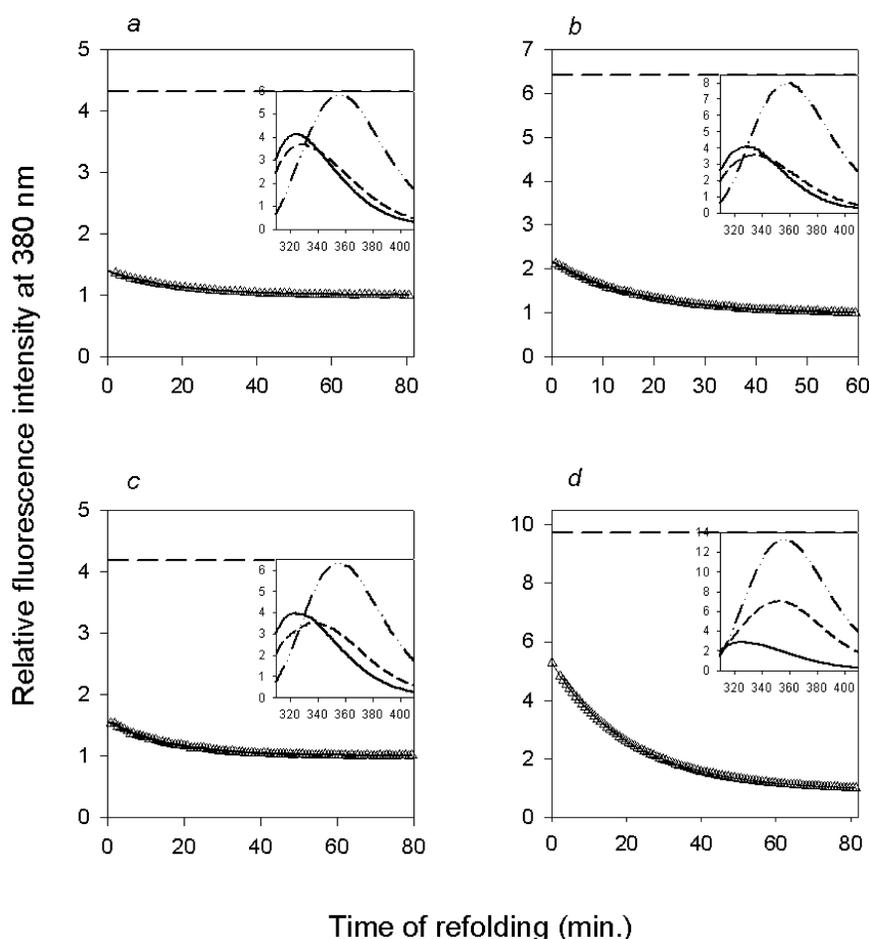


Figure 1. Slow refolding kinetics monitored by FRET. (a) Cys25TNB; (b) Cys40TNB; (c) Cys62TNB; and (d) Cys82TNB. Folding was monitored by the change in fluorescence intensity at 380 nm, upon excitation at 295 nm. Refolding was initiated by dilution of the protein in 6.1 M urea to 1.4 M urea. In all panels, the discontinuous lines represent the fluorescence signal of the unfolded protein. All data were normalized to a value of 1 for the fluorescence intensity of the native protein. The lines through the data were obtained by fitting to equation (1). The amplitude of the slow, observable kinetic phase, relative to the total change in FRET efficiency, upon folding from the unfolded to the native protein, is 12% for Cys25TNB, 21% for Cys40TNB, 17% for Cys62TNB and 49% for Cys82TNB. The inset in each panel shows the fluorescence spectrum of unfolded protein (— · —), native protein (—) and at one minute of refolding (— · —). For each protein, the spectra were normalized to a value of 1 for the intensity of the N state at 380 nm.

of unlabeled Cys25, Cys40, Cys62 and Cys82 in the native and unfolded forms are shown in Table 1. In the native state, all four proteins show a predominant fluorescence lifetime of ~ 4.9 ns, and a small fraction (~ 5 – 10%) of a faster decay component. The faster decay component could be due to a small fraction of Trp53 rotamers having faster de-excitation. Upon unfolding, three lifetimes are observed. The multi-exponential decays of tryptophan fluorescence have been attributed to ground-state heterogeneity arising from the rotation about the C^α – C^β bond (χ_1 rotamers).^{22,34}

Upon labeling the single Cys residue in each protein with TNB, the fluorescence intensity decay is different, depending on the position of the TNB group (Table 1). The lifetime of Trp53 in the N-state of each of the TNB-labeled proteins is shortened because of quenching by the TNB moiety.⁷ The extent of quenching of the fluorescence of Trp53 depends on how far it is separated from the TNB group in the different mutant

proteins, and it is more in the folded than in the unfolded forms, indicating that the mechanism of quenching is FRET from Trp53 to the TNB group.^{6,7} Indeed, FRET analysis indicates that the distances between Trp53 and the variously placed TNB groups, are as predicted from the NMR structure.³⁵ For Cys40TNB, Cys62TNB and Cys82TNB, a short lifetime component (<1 ns) is predominant ($>90\%$), indicating a uniformly quenched population of N and N-like molecules. The minor (5– 10%) long lifetime component could be due to a small fraction ($<5\%$) of protein molecules not being labeled, as well as being due to some labeled protein molecules with the Trp53 rotamer being in an unfavorable orientation for FRET. FRET is extremely sensitive to small changes in distance, and changes due to local structural fluctuations when some parts of the molecule are extended could also give rise to a small fraction of molecules with long fluorescence lifetimes.

The case of Cys25TNB is slightly different. Here,

Table 1. Fluorescence lifetimes (τ) and their amplitudes (α) obtained from the discrete analysis of the fluorescence intensity decays of the single-Trp single-Cys mutant forms of barstar and their TNB-labeled forms (A) in the native state and (B) in the unfolded form

Protein	τ_1 (ns) (α_1)	τ_2 (ns) (α_2)	τ_3 (ns) (α_3)	τ_m (ns) ^a
<i>A. Native (N) state</i>				
Cys25	4.95 ± 0.01 (0.89 ± 0.01)	1.45 ± 0.10 (0.11 ± 0.01)		4.58 ± 0.05
Cys40	4.99 ± 0.01 (0.88 ± 0.01)	1.86 ± 0.10 (0.12 ± 0.01)		4.63 ± 0.04
Cys62	4.78 ± 0.02 (0.92 ± 0.008)	1.55 ± 0.20 (0.08 ± 0.01)		4.54 ± 0.06
Cys82	4.88 ± 0.02 (0.94 ± 0.02)	1.31 ± 0.10 (0.06 ± 0.01)		4.65 ± 0.03
Cys25-TNB	4.87 ± 0.12 (0.09 ± 0.005)	1.32 ± 0.10 (0.27 ± 0.06)	0.55 ± 0.04 (0.64 ± 0.15)	1.13 ± 0.01
Cys40-TNB	5.02 ± 0.10 (0.02 ± 0.015)	1.64 ± 0.11 (0.03 ± 0.005)	0.45 ± 0.01 (0.95 ± 0.003)	0.59 ± 0.005
Cys62-TNB	4.77 ± 0.06 (0.07 ± 0.001)	1.08 ± 0.11 (0.04 ± 0.002)	0.10 ± 0.005 (0.89 ± 0.001)	0.45 ± 0.01
Cys82-TNB	4.73 ± 0.06 (0.02 ± 0.008)	1.30 ± 0.26 (0.03 ± 0.006)	0.24 ± 0.006 (0.95 ± 0.005)	0.38 ± 0.01
<i>B. Unfolded (U) form</i>				
Cys25	5.50 ± 0.25 (0.45 ± 0.05)	2.40 ± 0.40 (0.41 ± 0.01)	0.54 ± 0.10 (0.14 ± 0.05)	3.55 ± 0.04
Cys40	5.46 ± 0.10 (0.46 ± 0.10)	2.25 ± 0.31 (0.43 ± 0.02)	0.40 ± 0.11 (0.11 ± 0.06)	3.54 ± 0.03
Cys62	5.40 ± 0.19 (0.45 ± 0.05)	2.40 ± 0.42 (0.39 ± 0.01)	0.73 ± 0.09 (0.15 ± 0.05)	3.51 ± 0.06
Cys82	5.73 ± 0.21 (0.40 ± 0.06)	2.6 ± 0.35 (0.43 ± 0.01)	0.74 ± 0.25 (0.16 ± 0.06)	3.60 ± 0.12
Cys25-TNB	4.87 ± 0.23 (0.39 ± 0.01)	2.15 ± 0.16 (0.43 ± 0.01)	0.56 ± 0.11 (0.18 ± 0.01)	2.92 ± 0.13
Cys40-TNB	4.08 ± 0.23 (0.27 ± 0.03)	1.92 ± 0.18 (0.49 ± 0.02)	0.50 ± 0.07 (0.25 ± 0.02)	2.14 ± 0.07
Cys62-TNB	4.75 ± 0.13 (0.08 ± 0.01)	1.78 ± 0.03 (0.49 ± 0.006)	0.49 ± 0.01 (0.43 ± 0.02)	1.48 ± 0.08
Cys82-TNB	5.05 ± 0.07 (0.40 ± 0.03)	2.14 ± 0.06 (0.43 ± 0.01)	0.41 ± 0.11 (0.17 ± 0.03)	3.00 ± 0.08

^a Mean lifetime, $\tau_m = \sum \alpha_i \tau_i / \sum \alpha_i = 1$.

the two short lifetimes are very close in value (0.55 ns, 1.32 ns), and a clear distinction of their relative amplitudes is difficult (the error in the relative amplitudes of the two short lifetimes in Cys25TNB was ~20%, compared to the <5% error for the relative amplitudes in the other three proteins). For Cys25TNB, the combined amplitude of τ_1 and τ_2 appears to be the major component of the intensity decay, while for the other three labeled proteins, the intensity decay in the native protein is dominated by a single short lifetime component. It is possible that the Trp53–Cys25 distance in the N state is heterogeneous due to flexibility in the Cys25 region, and there could exist different populations with slightly different Trp53–Cys25 distances. Heteronuclear NMR³⁶ as well as hydrogen exchange experiments,³⁷ however, have not provided any evidence for flexibility in the Cys25 region.

A comparison of the mean lifetimes of the N state and U form indicates that energy transfer efficiency in the former (Table 1) is far more than that in the latter (Table 1), for each of the proteins. The FRET efficiency for the different FRET pairs was calculated from the predominant (>90% amplitude) fluorescence lifetime component in the

unlabeled and TNB-labeled proteins using equation (2). The Trp–CysTNB distances in the N state were then calculated from the FRET efficiency using previously determined values of R_0 for these four FRET pairs.⁶ The distances, shown in Table 2, are similar to those calculated from steady-state fluorescence intensity measurements for these FRET pairs.⁶ MEM and discrete analyses of the fluorescence intensity decays gave similar results.

Fluorescence lifetimes indicate structural heterogeneity in the I_L ensemble

Figure 2 shows the area-normalized fluorescence lifetime distributions obtained from the fluorescence intensity decays of all the four proteins in the U form, at one minute of refolding, and after completion of refolding in 1.4 M urea. The fluorescence lifetimes and their amplitudes obtained from the discrete analysis of the fluorescence intensity decays are given in Table 3. For all the proteins, the fluorescence lifetime distributions at one minute of refolding, especially for lifetimes below 1 ns (Figure 2 ii), do not deviate significantly from the N-state lifetime distributions (Figure 2 iii). Further, the long lifetimes (>1 ns) appear to be

Table 2. Energy transfer parameters in the native state

FRET pair	$J \times 10^{-13}$ (M ⁻¹ cm ⁻¹ nm ⁴)	R_0 (Å)	Lifetime (ns)		E	D–A distance (Å)
			(τ_d)	(τ_{da}) ^a		
Trp53–Cys25	7.4	26.9	4.95	0.78	0.84	19.0
Trp53–Cys40	5.1	25.2	4.99	0.45	0.91	17.1
Trp53–Cys62	7.3	26.8	4.78	0.10	0.98	14.1
Trp53–Cys82	7.7	27.0	4.88	0.24	0.95	16.5

^a τ_{da} corresponds to the major component (amplitude >0.9) obtained from the discrete analysis of the lifetime decays. It corresponds to τ_3 for Cys40, Cys62, Cys82 and an average of τ_2 and τ_3 for Cys25 (see the text).

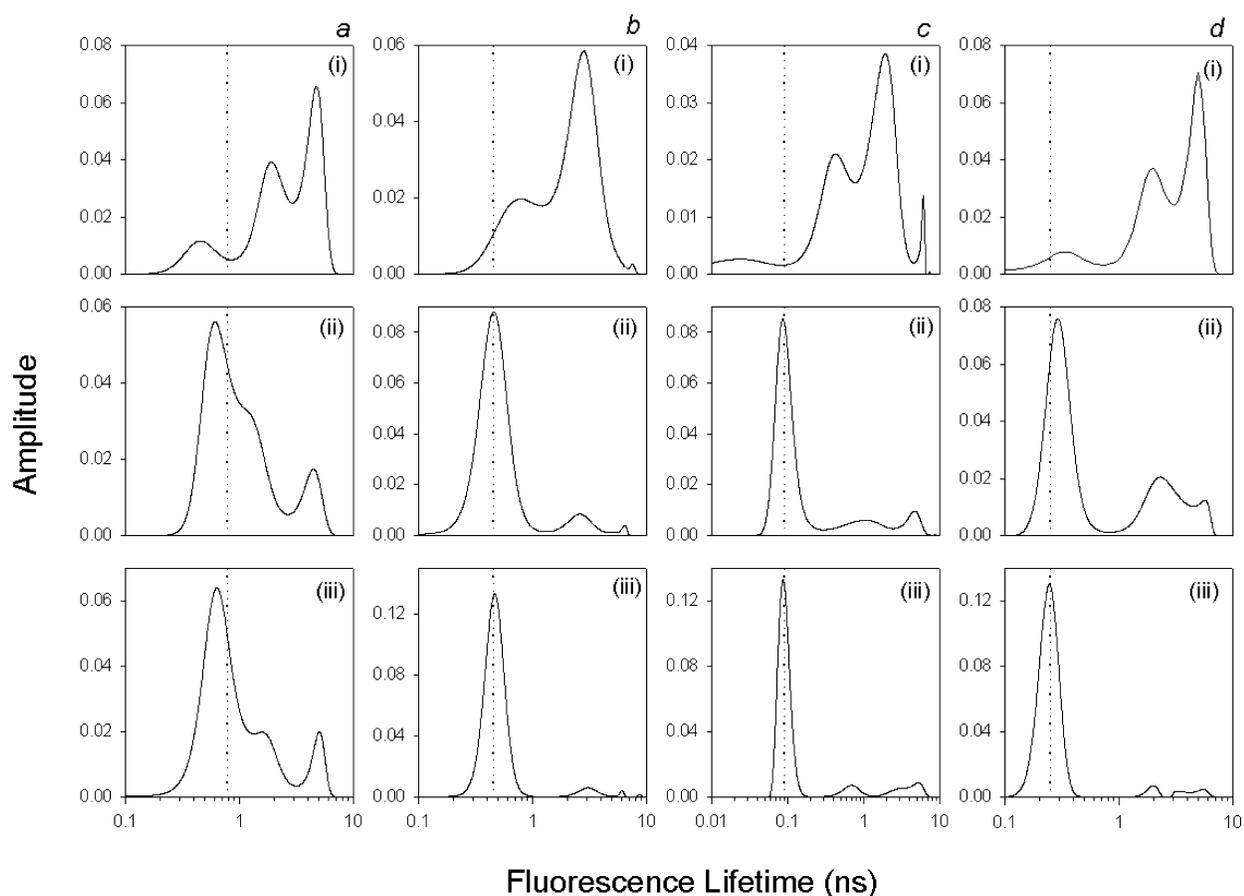


Figure 2. Fluorescence lifetime distributions of the TNB-labeled proteins. Lifetime distributions of (a) Cys25TNB, (b) Cys40TNB, (c) Cys62TNB and (d) Cys82TNB obtained by MEM analysis of the fluorescence intensity decays of Trp53 in the unfolded form in 6.1 M urea (panels i), at one minute of refolding in 1.4 M urea (panels ii) and after completion of refolding in 1.4 M urea (panels iii). The dotted lines in all the panels represent the peaks of the lifetime distributions in the native state for (b), (c) and (d). In (a), a mean lifetime for the broad N-like distribution is depicted by the dotted line (see the text).

arising from molecules having U-like distances. It is important to note the absence, at one minute of refolding, of any significant region in the fluorescence lifetime distribution different from the lifetime distributions obtained for either N or U molecules. This indicates the absence of intermediate states with intra-molecular distances in between those of the N and U forms, at one minute of refolding.

The refolding condition was altered by changing the final concentration of urea at which each protein was refolded. Concentrations of urea greater than 2 M were not used, because the

unfolding transition of N for each protein begins at this concentration of urea.⁶ At all concentrations of urea, and for all four proteins, only two distinct fluorescence lifetime distributions were observed at one minute of refolding. Distributions centered within ± 0.1 ns of the center of the distribution of the fully folded N-state are classified as N-like, while those centered at lifetimes greater than 1.5 ns are classified as U-like, keeping in mind the broad distribution of lifetimes in the U form. For each of the four FRET pairs, the N-like lifetime arises from molecules possessing an N-like separation and the U-like lifetime arises from molecules

Table 3. Fluorescence lifetimes (τ) and their relative amplitudes (α) obtained from the discrete analysis of the fluorescence intensity decays collected at 1 minute of refolding upon dilution of urea from 6.1 M to 1.4 M

Protein	Fluorescence lifetimes			
	τ_1 (ns) (α_1)	τ_2 (ns) (α_2)	τ_3 (ns) (α_3)	τ_m (ns)
Cys25-TNB	4.60 ± 0.17 (0.10 ± 0.01)	1.38 ± 0.20 (0.37 ± 0.10)	0.57 ± 0.06 (0.53 ± 0.20)	1.27 ± 0.01
Cys40-TNB	4.37 ± 0.17 (0.04 ± 0.008)	1.48 ± 0.15 (0.10 ± 0.01)	0.44 ± 0.01 (0.86 ± 0.01)	0.69 ± 0.01
Cys62-TNB	4.59 ± 0.08 (0.07 ± 0.005)	1.16 ± 0.12 (0.11 ± 0.005)	0.11 ± 0.01 (0.81 ± 0.006)	0.56 ± 0.01
Cys82-TNB	4.73 ± 0.06 (0.13 ± 0.01)	1.91 ± 0.19 (0.22 ± 0.01)	0.29 ± 0.006 (0.64 ± 0.01)	1.24 ± 0.01

possessing a U-like separation, between the donor and acceptor fluorophores.⁷ Since the N-like distribution is narrow for each of the four proteins, it is expected that if any partially folded form, with a CysTNB–Trp53 separation intermediate between that in the N and U forms, were to be populated to any significant level, it should be detected. For Cys25TNB, Cys40TNB and Cys62TNB, the short lifetime distribution, at one minute of refolding in 1.4 M urea, is virtually indistinguishable from the lifetime distribution of the N-state (Figure 2). As a result, the forms of the protein present at one minute of folding cannot be distinguished from the N-state on the basis of the Trp53–Cys25TNB, Trp53–Cys40TNB and Trp53–Cys62TNB distances. For Cys82TNB, on the other hand, the center of the short lifetime distribution is shifted to a lifetime longer than that of the N-state (Figure 2), indicating clearly that all molecules displaying an N-like fluorescence lifetime distribution at one minute of refolding are not present as N (see below).

Only the relative amplitudes of the N-like and U-like lifetimes were found to vary with the final concentration of urea. The relative amplitude corresponding to each lifetime reflects the fraction of molecules having a particular separation between donor and acceptor fluorophores. The relative amplitude of the N-like lifetime in different urea concentrations is shown for all the proteins in Figure 3. The amplitudes of the short lifetimes of the N state at the same concentrations of urea are shown not to change significantly in the range of concentrations used, confirming that native protein remains fully folded. It should be noted that the four intra-molecular distances that have been measured in the TRFRET studies reported here represent distances from the core (Trp53)³⁵ to points in four different secondary structural elements of the protein: Cys25 in helix 1, Cys40 in helix 2, Cys62 in helix 3 and Cys82 at the beginning of the third strand of the three-stranded β -sheet. Two of the distances (Trp53–Cys82TNB and Trp53–Cys40TNB) represent distances within the extended core, while two of the distances (Trp53–Cys25TNB and Trp53–Cys62TNB) represent distances from the core to the surface.

In the case of each TNB-labeled protein, the relative amplitude of the short lifetime component at one minute of refolding corresponds to the fraction of molecules having a native-like separation between the donor and acceptor fluorophores comprising the FRET pair. This fraction naturally includes those molecules present as N, and double-jump experiments³² with Cys40 and Cys82, indicate that 40% of molecules are present as N at one minute of refolding (see above). Since the remaining 60% molecules are present either as I_L or as U, the change in the relative amplitude of the short N-like lifetime with urea concentration (Figure 3) indicates the change in the fraction of I_L molecules possessing an N-like separation. The observation that the fluorescence lifetimes of I_L molecules are always either N-like or U-like indi-

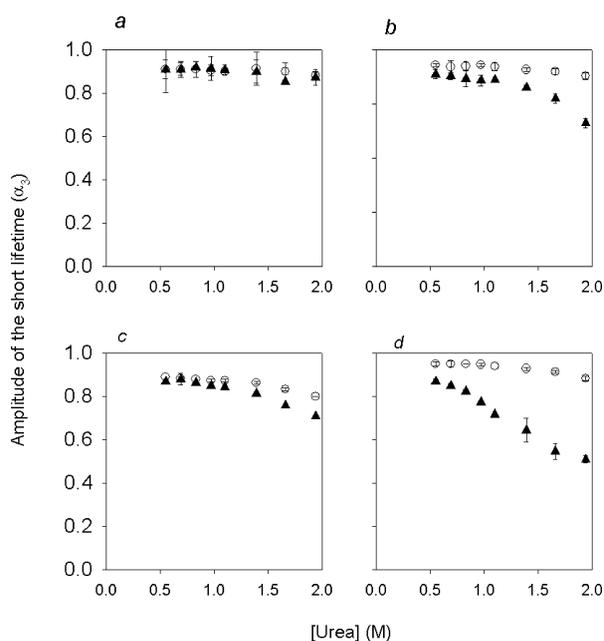


Figure 3. Urea-dependence of heterogeneity in the I_L ensemble. The fractional amplitude of the shortest lifetime component obtained from the discrete as well as MEM analysis of the fluorescence intensity decays of (a) Cys25TNB, (b) Cys40TNB, (c) Cys62TNB and (d) Cys82TNB, is shown at different urea concentrations. Open circles correspond to the protein at equilibrium, triangles correspond to the protein at one minute of refolding. Refolding was initiated by diluting the urea concentration from 6.1 M to different concentrations below 2 M urea, shown here. The error bars represent standard deviations from three repetitions of each fluorescence decay measurement. For many data points, the error bar is smaller than the size of the symbol.

cates that the four distances in I_L are either N-like or U-like at all urea concentrations.

The TRFRET data in Figure 3 indicate that I_L is best described as an ensemble of conformations that are structurally distinct, and show how four intra-molecular distances vary within the I_L ensemble, with a change in conditions that confer different stabilities. In strongly stabilizing conditions (0.5 M urea), the four distances in virtually all I_L molecules are the same as they are in N (Figure 3). Thus, nearly all I_L molecules appear to be as specifically and uniformly collapsed as N, in strongly stabilizing conditions. In contrast, in less stabilizing conditions (1.4 M urea), 99% of the molecules have an N-like Trp53–Cys25TNB distance, 93% of the molecules have an N-like Trp53–Cys40TNB distance, 95% of the molecules have an N-like Trp53–Cys62TNB distance, but only 71% of the molecules have an N-like Trp53–Cys82TNB distance. In other words, there appear to be, in marginally stabilizing (but native-like) conditions, different populations of molecules in each of which one or more distances are unfolded-like. Since the N states of the four TNB-labeled proteins have similar stabilities,⁶ the differences in the denaturant-dependences of the relative amplitudes

of the short lifetime for the four labeled proteins (Figure 3) must reflect the differences in the stabilities of different structural components comprising the I_L ensemble.

Folding of I_L monitored by fluorescence lifetime changes in Cys82TNB

Figure 4a shows the fluorescence lifetime distributions obtained by MEM analysis of the fluorescence intensity decays collected at different times during the slow folding reaction of Cys82TNB in 1.4 M urea. The distribution at 0.8 minute of refolding indicates that $\sim 60\%$ of the molecules are as compact as the N-state with respect to the Trp53–Cys82TNB distance. Since $\sim 40\%$ molecules are present as N and $\sim 60\%$ as I_L (see above) at this time of refolding in 1.4 M urea,³² the characteristics of the 60% molecules present as I_L can be determined after subtracting out the contribution of the 40% molecules that have already folded to N. It then appears that

with respect to the Trp53–Cys82TNB distance, one-third of the I_L molecules ($\sim 20\%$ molecules of all protein molecules) are as compact as the N state, and two-thirds of the I_L molecules ($\sim 40\%$ of all molecules) are as expanded as the U form. The I_L molecules with an N-like Trp53–Cys82TNB distance at one minute are not native, but still need to undergo minor structural readjustment: the discrete value of the short fluorescence lifetime, as also the peak of the MEM distribution, changes from $0.3(\pm 0.006)$ ns at one minute of refolding to $0.25(\pm 0.006)$ ns at the end of the folding reaction (Figure 4a). The small difference in the lifetimes of I_L and N, corresponds to very small changes in distance. The change in the short lifetime from 0.3 to 0.25 ns was obtained reproducibly in three repetitions of the experiment, when the data were subjected to either MEM or discrete analysis, confirming the exquisite sensitivity of the TRFRET method for detecting even small changes in structure.⁷

The small but significant shift in the position of

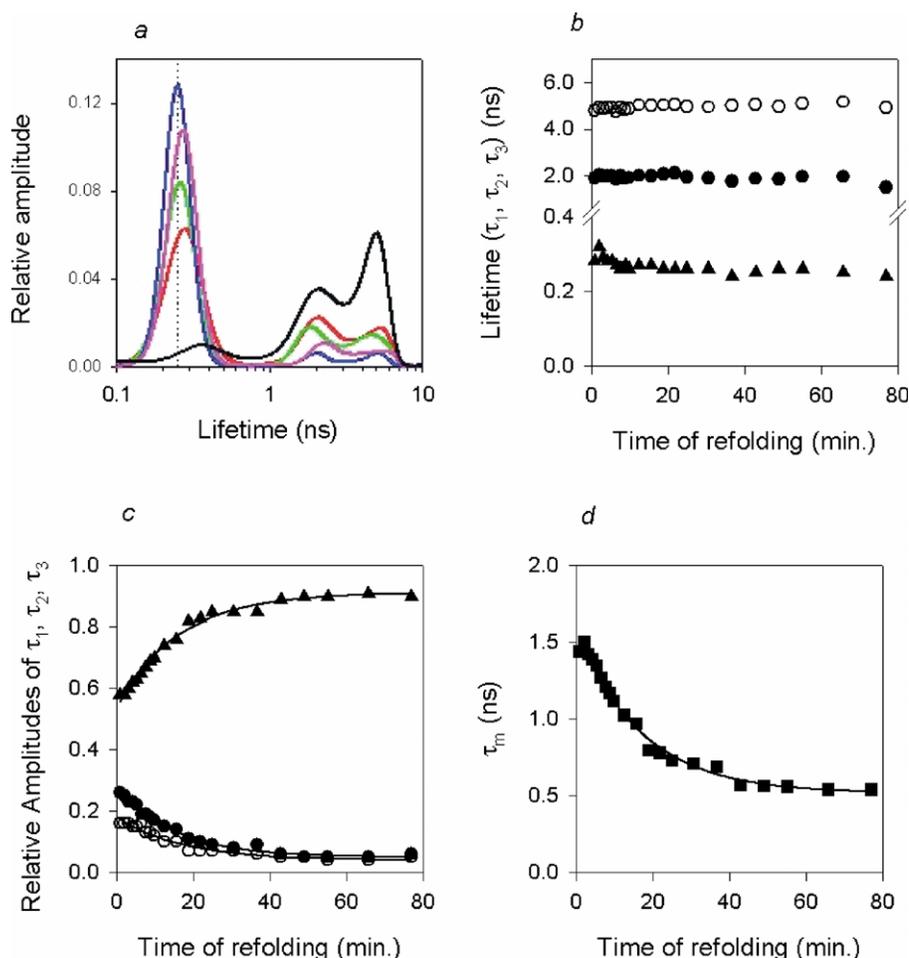


Figure 4. TRFRET-monitored kinetics of refolding of Cys82TNB in 1.4 M urea. (a) MEM distributions of the unfolded protein in 6.1 M urea (black), at 0.8 minute after refolding (red), at 9.8 minutes after refolding (green), at 24.9 minutes after refolding (pink) and after complete refolding in 1.4 M urea (blue). The dotted line indicates the peak position of the lifetime distribution in the N state. (b) The lifetimes obtained from the discrete analysis of the fluorescence intensity decays during refolding, τ_1 (○), τ_2 (●), τ_3 (▲). (c) Relative amplitudes (α_i) of τ_1 (○), τ_2 (●), τ_3 (▲) ($\sum \alpha_i = 1$). (d) The mean fluorescence lifetime ($\tau_m = \sum \alpha_i \tau_i$) during refolding. The continuous lines in (c) and (d) are fits to equation (1).

the distribution of fluorescence lifetimes as folding proceeds (Figure 4a), suggests that the Trp53–Cys82TNB distance decreases continuously, and not in discrete steps, even at late times of folding. A similar continuous decrease is not observed for the three other distances being measured (Figure 2), at least at late times of folding. It remains to be seen if these distances change continuously at early times of folding; these measurements are now in progress. In this context, it should be noted that a TRFRET characterization of the equilibrium unfolding of barstar⁷ had suggested previously that the unfolding transition is continuous, and does not occur in a step-wise manner.

The changes in the lifetimes and their relative amplitudes during refolding are shown in Figure 4b and c. The discrete analysis of the same data agrees very well with the peak positions and relative amplitudes obtained by the MEM analysis. The amplitude of the shortest lifetime component increases at a rate of 0.06 min^{-1} , which is the same as the rate of the decrease in the amplitude of either of the two longer lifetimes. This rate is also the same as the apparent rate of folding monitored by the changes in steady-state FRET efficiency during folding, for all the proteins (Figure 1). It agrees well with the rate of overall compaction from I_L to the N state, as measured earlier by time-resolved fluorescence anisotropy decay measurements.²⁰

Discussion

Analysis of the FRET between the core Trp residue of barstar with Cys residues at four different positions reveals the structural features of the I_L ensemble. It appears that in marginally stable conditions, the C-terminal region, which comprises the third strand of the three-stranded β -sheet of barstar, is unstructured in most members of the I_L ensemble. In contrast, in very stable conditions (0.5 M urea), I_L possesses uniformly native-like compactness (see above). These results support an important implicit prediction of energy landscape theory that different folding pathways will appear to dominate under different folding conditions. In very stable conditions, the dominant folding pathway is one in which uniformly native-like compactness is achieved during the formation of I_L . On the other hand, in marginally stable conditions, the dominant folding pathway is one in which uniformly native-like compactness is achieved only later, during subsequent folding of I_L .

A critical issue in protein folding studies concerns the nature of the initial hydrophobic collapse of the polypeptide chain, which has been observed for many proteins, in experiments^{23,30,38} as well as in lattice model simulations.³⁹ While several studies have suggested that the initial collapse leads to the formation of a specific intermediate,^{40,41}

other studies have suggested that the collapse merely represents a non-specific, solvent-dependent chain contraction 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.^{42–44} The analysis of the structure of I_L , presented here, can be used to make inferences concerning the nature of the collapse leading to the formation of I_E , which precedes I_L on the folding pathway, because folding reactions are most likely to be hierarchal, in that structure accumulates progressively: it is unlikely that if parts of the chain were collapsed in I_E , they would open out in I_L . If the initial collapse reaction to I_E were non-specific, it is unlikely that a segment of the polypeptide chain would remain as extended in I_L as it is in U, as is observed. This observation indicates that the initial collapse to I_E itself must have led to specific compaction of only some segments of the polypeptide chain. Such an argument gains credence from noting that the distance Trp53–Cys82TNB, which is observed to be U-like in I_L , is a distance separating two points in the hydrophobic core of the protein, and hence would be expected to have contracted in any initial non-specific collapse.

In single-site TRFRET studies of the refolding of cytochrome *c*,^{3,23} FRET between the heme group and a dansyl group attached to the C-terminal Cys residue had indicated a distribution of N-like and U-like populations throughout refolding. But it is possible that the apparently U-like molecules are actually compact, with only the one distance that was being monitored, being U-like. For example, if only the Trp53–Cys82TNB FRET pair had been monitored in the case of barstar, it would have appeared that a large fraction of the molecules is completely unfolded at one minute of refolding. Monitoring four distances has shown clearly that this is not the case. This study therefore highlights the importance of using multiple FRET pairs, because information obtained using only one distance may be misleading. Of course, the use of more FRET pairs increases the structural resolution obtained.

Implications of heterogeneity in a late protein folding ensemble

The data in Figures 3 and 4 show that as the overall stability of I_L is decreased by an increase in concentration of denaturant, different structural forms in the I_L ensemble are destabilized differentially so that structural heterogeneity increases. The increase in structural heterogeneity makes it appear that some parts of the structure (measured by the four TRFRET-monitored intra-molecular distances) are lost at lower concentration of urea than others. In very stable conditions, those conformations of the I_L ensemble are most stabilized, in which all four distances monitored, are N-like. In less stable conditions, the I_L ensemble has much

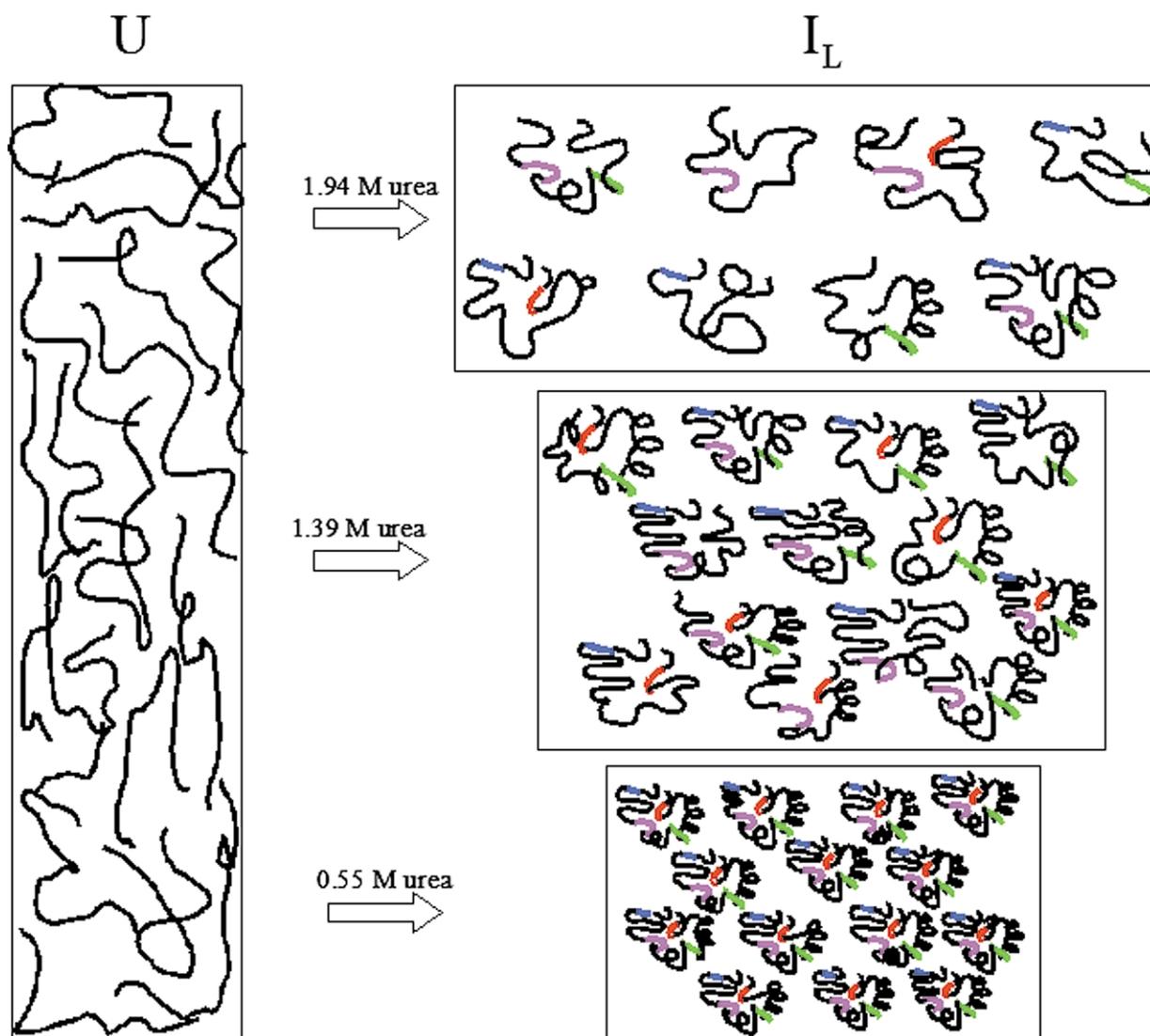


Figure 5. Structural heterogeneity in I_L . The cartoon depicts the structural implications of the data shown in Figure 3. The highly heterogeneous U ensemble folds to a different I_L ensemble at different urea concentrations. Formation of a native-like intra-molecular distance in I_L is depicted by a colored line. Four intra-molecular distances in barstar were monitored here; thus, only four colored lines are shown. During refolding in a low concentration of urea (0.55 M), almost all the molecules have all four distances similar to those in the native state. At higher urea concentrations, either one, two, three or four distances are native-like.

greater structural heterogeneity, with significant populations of molecules with one or more of the four distances being U-like. In this way, the structural properties of I_L depend on the nature of the refolding conditions, as depicted in Figure 5.

It is important to understand why structural heterogeneity persists in a late intermediate ensemble such as I_L . Structural heterogeneity in I_L implies that I_L possesses high configurational entropy. Earlier time-resolved fluorescence anisotropy decay experiments²⁰ had, in fact, suggested this: Trp53 in the core of I_L was observed to undergo substantial, albeit partially hindered, rotational motion, indicating substantial configurational flexibility even in the core. The rate limiting step in the folding of barstar has been shown to involve consolidation of the core,⁴⁵ which occurs as the first event during the non-cooperative

$I_L \rightarrow N$ reaction.²⁰ The transition state is expected to resemble the intermediate nearest to it in energy, which is usually the one that immediately precedes it on the reaction pathway. The observation that I_L possesses and is probably stabilized by high configurational entropy would therefore indicate that the transition state, which follows I_L , also possesses and is stabilized by high configurational entropy. This offers an explanation for the structural heterogeneity seen in I_L : it contributes to stabilization of the transition state that follows, and folding is speeded. Lattice model simulations indicate that a transition state with many configurations leads rapidly to the native state, providing a solution to the Levinthal paradox.^{39,46}

The existence of structural heterogeneity in a late intermediate ensemble implies that similar or greater heterogeneity must exist in any earlier

intermediate ensemble. In fact, heterogeneity in the initial intermediate ensemble, which accumulates within a few milliseconds or faster, was detected first in the case of barstar,²⁸ and subsequently, for ribonuclease A,^{47,48} lysozyme,⁴⁹ cytochrome *c*,³⁸ and apomyoglobin.⁵⁰ In the case of thioredoxin, the observation that the denaturant-dependence of the amplitude of the initial (few milliseconds) kinetic phase is different when monitored by circular dichroism and fluorescence,⁵¹ can be interpreted to be representing structural heterogeneity in the initial intermediate ensemble.

The structural transition that is being monitored during the slow folding reaction of I_L is coupled to *trans* to *cis* isomerization of the Tyr47-Pro48 peptide bond.^{27,29} It is, however, very unlikely that proline isomerization is directly responsible for the structural heterogeneity that has been identified in this study, because this heterogeneity is dependent on the concentration of urea, whereas proline isomerization is known not to be affected by denaturant. It appears that multiple parallel folding reactions, known to occur during the folding of barstar, are responsible for the observed heterogeneity.^{28,31} In fact, there is strong evidence, albeit not as direct as the TRFRET data presented here for I_L, showing that different components of the early folding intermediate ensemble, I_E, of barstar are stabilized differentially under conditions that confer different stability.^{28,41}

According to the energy landscape view of protein folding, heterogeneity in a folding intermediate ensemble is a direct consequence of multiple folding routes.^{1,2,39} Several earlier experimental studies have, in fact, shown that multiple intermediates and multiple folding pathways exist for a number of proteins, including barstar,^{28,31} but there has been no experimental evidence showing that increased stability leads to a reduction in structural heterogeneity of any folding intermediate ensemble. Energy landscape theory suggests that the reason for the bias in the energy surface to favor the native state may be an energy gradient and/or a restriction in accessible conformational space.³⁹ Here, the use of the TRFRET methodology to distinguish between different structural forms comprising an intermediate ensemble has shown clearly that increasing stability reduces conformational heterogeneity in the late folding intermediate ensemble of barstar. The fundamental result reported here, that the structure apparent in a folding intermediate depends on the conditions employed to study folding, implies that the folding pathway observed for a given protein will appear different under different conditions.

Materials and Methods

The purification and characterization of the four proteins Cys25, Cys40, Cys62 and Cys82 has been described.⁶ All experiments were done at 10 °C, in buffer

containing 20 mM Tris and 0.25 mM EDTA at pH 8.0. The extent of TNB labeling for all the proteins was confirmed to be >98% by mass spectrometry. The final protein concentrations used in folding experiments were 20–40 μM.

The folding reactions were initiated by manual mixing of the required solutions in the fluorescence cuvette. The protein was unfolded by incubating in 6.1 M urea-containing buffer, for three hours. Folding was initiated by diluting the unfolded protein solution to obtain the desired urea concentration. A SPEX Fluorolog FLIT11 fluorimeter was used for steady-state fluorescence measurements. Tryptophan fluorescence excited at 295 nm, was monitored at 380 nm. The dead-time of measurement was ~10 s. Time-resolved fluorescence decay curves were collected as described.⁵² The dead-time of the measurements was ~15 s and each decay was collected for ~45 s.

Analysis of steady-state fluorescence monitored kinetics

The change in fluorescence intensity (*y*) observed for the refolding reactions as a function of time, *t*, was fit to a single-exponential equation, given by:

$$y = y_0 + a e^{-\lambda t} \quad (1)$$

Analysis of time-resolved fluorescence intensity decays

Analyses of fluorescence intensity decays were carried out by discrete analysis²⁰ as well as by MEM analysis.^{21,22,52–54}

MEM analysis begins with the assumption that the decay originates from a distribution of fluorescence lifetimes in the range 10 ps to 10 ns, or in a similar range, with all lifetimes having equal probability (amplitude). Subsequently, the distribution is modified in each iteration of the analysis, leading to minimization of the residuals (χ^2) and maximization of the Shannon–Jaynes entropy, $S = -\sum P_i \log P_i$, where P_i is the probability (amplitude) of the *i*th lifetime. For a particular value of χ^2 , there could be many possible values of P_i . MEM analysis identifies the distribution for which S is maximum. The analysis is terminated when successive iterations do not change the values of χ^2 , S , and the distribution profile. Thus, MEM analysis is independent of any physical model or mathematical equation (such as Gaussian or Lorentzian distributions), and will not yield any feature in the distribution of lifetimes, unless warranted by the data.

The robustness of the lifetime distributions⁷ (including peak positions and widths of distributions) obtained by MEM analysis was checked exhaustively by collecting data on several samples under the same sample conditions. Peak values of MEM distributions agreed with those obtained from discrete lifetime analyses, within ~5%.

Analysis of FRET

The efficiency of energy transfer for a donor–acceptor pair depends on the distance *R*, between donor and acceptor, and can be determined experimentally from the fluorescence lifetime of the donor–acceptor samples

(τ_{DA}) and of the donor-alone sample (τ_D):

$$E = \left[1 + \frac{R^6}{R_0^6} \right]^{-1} = 1 - \frac{\tau_{DA}}{\tau_D} \quad (2)$$

where R_0 is the Forster's distance. The calculation of R_0 for all the FRET pairs has been described.⁶

Acknowledgements

We thank M. K. Mathew for discussion and useful insights into the interpretation of the data, and A.S.R. Koti for discussion and technical assistance. K.S. is the recipient of a Kanwal Rekhi Scholarship for Career Development. J.B.U. was the recipient of a Swarnajayanti Fellowship from the Government of India. This work was funded by the Tata Institute of Fundamental Research, and by the Department of Science and Technology, Government of India.

References

- Dill, K. A. & Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nature Struct. Biol.* **4**, 10–19.
- Onuchic, J. N., Luthey-Schulten, Z. & Wolynes, P. G. (1997). Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* **48**, 545–600.
- Tezcan, F. A. *et al.* (2002). Using deeply trapped intermediates to map the cytochrome *c* folding landscape. *Proc. Natl Acad. Sci. USA*, **99**, 8626–8630.
- Lee, J. C., Engman, K. C., Tezcan, F. A., Gray, H. B. & Winkler, J. R. (2002). Structural features of cytochrome *c'* folding intermediates revealed by fluorescence energy-transfer kinetics. *Proc. Natl Acad. Sci. USA*, **99**, 14778–14782.
- Teilum, K., Maki, K., Kragelund, B. B., Poulsen, F. M. & Roder, H. (2002). Early kinetic intermediate in the folding of acyl-CoA binding protein detected by fluorescence labeling and ultrarapid mixing. *Proc. Natl Acad. Sci. USA*, **99**, 9807–9812.
- Sridevi, K. & Udgaonkar, J. B. (2003). Surface expansion is independent of and occurs faster than core solvation during the unfolding of barstar. *Biochemistry*, **42**, 1551–1563.
- 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.
- Wu, P. & Brand, L. (1994). Resonance energy transfer: methods and applications. *Anal. Biochem.* **218**, 1–13.
- Lillo, M. P., Beechem, J. M., Szpikowska, B. K., Sherman, M. A. & Mas, M. T. (1994). Design and characterization of a multisite fluorescence energy-transfer system for protein folding studies: a steady-state and time-resolved study of yeast phosphoglycerate kinase. *Biochemistry*, **36**, 11261–11272.
- dos Remedios, C. G., Miki, M. & Barden, J. A. (1987). Fluorescence resonance energy transfer measurements of distances in actin and myosin. A critical evaluation. *J. Muscle Res. Cell. Motil.* **8**, 97–117.
- Navon, A., Ittah, V., Landsman, P., Scheraga, H. A. & Haas, E. (2001). Distributions of intra-molecular distances in the reduced and denatured states of bovine pancreatic ribonuclease A. Folding initiation structures in the C-terminal portions of the reduced protein. *Biochemistry*, **40**, 105–118.
- 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.
- Arai, M., Ito, K., Inobe, T., Nakao, M., Maki, K., Kamagata, K. *et al.* (2002). Fast compaction of alpha-lactalbumin during folding studied by stopped-flow X-ray scattering. *J. Mol. Biol.* **321**, 121–132.
- Pollack, L., Tate, M. W., Finnefrock, A. C., Kalidas, C., Trotter, S., Darnton, N. C. *et al.* (2001). Time resolved collapse of a folding protein observed with small angle X-ray scattering. *Phys. Rev. Lett.* **86**, 4962–4965.
- Noppert, A., Gast, K., Zirwer, D. & Damaschun, G. (1998). Initial hydrophobic collapse is not necessary for folding RNase A. *Fold. Des.* **3**, 213–221.
- Gast, K., Zirwer, D., Muller-Frohne, M. & Damaschun, G. (1998). Compactness of the kinetic molten globule of bovine alpha-lactalbumin: a dynamic light scattering study. *Protein Sci.* **7**, 2004–2011.
- Jones, B. E., Beechem, J. M. & Matthews, C. R. (1995). Local and global dynamics during the folding of *Escherichia coli* dihydrofolate reductase by time-resolved fluorescence spectroscopy. *Biochemistry*, **34**, 1867–1877.
- Canet, D., Doering, K., Dobson, C. M. & Dupont, Y. (2001). High-sensitivity fluorescence anisotropy detection of protein-folding events: application to alpha-lactalbumin. *Biophys. J.* **80**, 1996–2003.
- Bilsel, O., Yang, L., Zitzewitz, J. A., Beechem, J. M. & Matthews, C. R. (1999). Time-resolved fluorescence anisotropy study of the refolding reaction of the alpha-subunit of tryptophan synthase reveals non-monotonic behavior of the rotational correlation time. *Biochemistry*, **38**, 4177–4187.
- 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.
- Brochon, J. C. (1994). Maximum entropy method of data analysis in time-resolved spectroscopy. *Methods Enzymol.* **240**, 262–311.
- Swaminathan, R., Krishnamoorthy, G. & Periasamy, N. (1994). Similarity of fluorescence lifetime distributions for single tryptophan proteins in the random coil state. *Biophys. J.* **67**, 2013–2023.
- Lyubovitsky, J. G., Gray, H. B. & Winkler, J. R. (2002). Mapping the cytochrome *c* folding landscape. *J. Am. Chem. Soc.* **124**, 5481–5485.
- Deniz, A. A., Laurence, T. A., Beligere, G. S., Dahan, M., Martin, A. B., Chemla, D. S. *et al.* (2000). Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proc. Natl Acad. Sci. USA*, **97**, 5179–5184.
- Schuler, B., Lipman, E. A. & Eaton, W. A. (2002). Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature*, **419**, 743–747.
- Rhoades, E., Gussakovsky, E. & Haran, G. (2003).

- Watching proteins fold one molecule at a time. *Proc. Natl Acad. Sci. USA*, **100**, 3197–3202.
27. Schreiber, G. & Fersht, A. R. (1993). The refolding of *cis*- and *trans*-peptidylprolyl isomers of barstar. *Biochemistry*, **32**, 11195–11203.
 28. Shastry, M. C. & Udgaonkar, J. B. (1995). The folding mechanism of barstar: evidence for multiple pathways and multiple intermediates. *J. Mol. Biol.* **247**, 1013–1027.
 29. Shastry, M. C., 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.
 30. Agashe, V. R., Shastry, M. C. & Udgaonkar, J. B. (1995). Initial hydrophobic collapse in the folding of barstar. *Nature*, **377**, 754–757.
 31. Bhuyan, A. K. & Udgaonkar, J. B. (1999). Observation of multistate kinetics during the slow folding and unfolding of barstar. *Biochemistry*, **38**, 9158–9168.
 32. Sridevi, K. (2003). Structural and temporal characterization of the folding pathway of barstar. PhD Thesis, National Centre for Biological Sciences, TIFR, Bangalore, India.
 33. 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.
 34. Colucci, W. J., Tilstra, L., Sattler, M. C., Fronczek, F. R. & Barkley, M. D. (1990). Conformational studies of a constrained tryptophan derivative: implications for the fluorescence quenching mechanism. *J. Am. Chem. Soc.* **112**, 9182–9190.
 35. Lubienski, M. J., Bycroft, M., Freund, S. M. & Fersht, A. R. (1994). Three-dimensional solution structure and ¹³C assignments of barstar using nuclear magnetic resonance spectroscopy. *Biochemistry*, **33**, 8866–8877.
 36. Sahu, S. C., Bhuyan, A. K., Majumdar, A. & Udgaonkar, J. B. (2000). Backbone dynamics of barstar: a (¹⁵N) NMR relaxation study. *Proteins: Struct. Funct. Genet.* **41**, 460–474.
 37. Bhuyan, A. K. & Udgaonkar, J. B. (1998). Two structural subdomains of barstar detected by rapid mixing NMR measurement of amide hydrogen exchange. *Proteins: Struct. Funct. Genet.* **30**, 295–308.
 38. Akiyama, S., Takahashi, S., Ishimori, K. & Morishima, I. (2000). Stepwise formation of alpha-helices during cytochrome c folding. *Nature Struct. Biol.* **7**, 514–520.
 39. Dinner, A. R., Sali, A., Smith, 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.
 40. Shastry, M. C. & Roder, H. (1998). Evidence for barrier-limited protein folding kinetics on the microsecond time scale. *Nature Struct. Biol.* **5**, 385–392.
 41. 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.
 42. Sosnick, T. R., Shtilerman, M. D., Mayne, L. & Englander, S. W. (1997). Ultrafast signals in protein folding and the polypeptide contracted state. *Proc. Natl Acad. Sci. USA*, **94**, 8545–8550.
 43. 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.
 44. Krantz, B. A., Mayne, L., Rumbley, 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.
 45. Rami, B. R. & Udgaonkar, J. B. (2002). Mechanism of formation of a productive molten globule form of barstar. *Biochemistry*, **41**, 1710–1716.
 46. Dobson, C. M., Sali, A. & Karplus, M. (1998). Protein folding: a perspective from theory and experiment. *Angew. Chem. Int. Ed.* **37**, 868–893.
 47. Houry, W. A., Rothwarf, D. M. & Scheraga, H. A. (1996). Circular dichroism evidence for the presence of burst-phase intermediates on the conformational folding pathway of ribonuclease A. *Biochemistry*, **35**, 10125–10133.
 48. Houry, W. A. & Scheraga, H. A. (1996). Structure of a hydrophobically collapsed intermediate on the conformational folding pathway of ribonuclease A probed by hydrogen–deuterium exchange. *Biochemistry*, **35**, 11734–11746.
 49. Morgan, C. J., Miranker, A. & Dobson, C. M. (1998). Characterization of collapsed states in the early stages of the refolding of hen lysozyme. *Biochemistry*, **37**, 8473–8480.
 50. Nishimura, C., Dyson, H. J. & Wright, P. E. (2002). The apomyoglobin folding pathway revisited: structural heterogeneity in the kinetic burst phase intermediate. *J. Mol. Biol.* **322**, 483–489.
 51. Georgescu, R. E., Li, J. H., Goldberg, M. E., Tasayco, M. L. & Chaffotte, A. F. (1998). Proline isomerization-independent accumulation of an early intermediate and heterogeneity of the folding pathways of a mixed alpha/beta protein, *Escherichia coli* thioredoxin. *Biochemistry*, **37**, 10286–10297.
 52. Lakshmikanth, G. S. & Krishnamoorthy, G. (1999). Solvent-exposed tryptophans probe the dynamics at protein surfaces. *Biophys. J.* **77**, 1100–1106.
 53. Skilling, J. & Bryan, R. K. (1984). Maximum entropy image reconstruction: general algorithm. *Mon. Not. R. Astr. Soc.* **211**, 111–124.
 54. Swaminathan, R. & Periasamy, N. (1996). Analysis of fluorescence decay by the maximum entropy method: Influence of noise and analysis parameters on the width of the distribution of lifetimes. *Proc. Indian Acad. Sci. Chem. Sci.* **108**, 39–49.

Edited by C. R. Matthews

(Received 12 August 2003; received in revised form 2 December 2003; accepted 9 December 2003)