

Characterization of Intra-molecular Distances and Site-specific Dynamics in Chemically Unfolded Barstar: Evidence for Denaturant-dependent Non-random Structure

Anoop M. Saxena¹, Jayant B. Udgaonkar^{2*} and G. Krishnamoorthy^{1*}

¹Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005 India

²National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bangalore 565 065 India

The structure and dynamics of the unfolded form of a protein are expected to play critical roles in determining folding pathways. In this study, the urea and guanidine hydrochloride (GdnHCl)-unfolded forms of the small protein barstar were explored by time-resolved fluorescence techniques. Barstar was labeled specifically with thionitrobenzoate (TNB), by coupling it to the thiol side-chain of a cysteine residue at one of the following positions on the sequence: 14, 25, 40, 42, 62, 82 and 89, in single cysteine-containing mutant proteins. Seven intra-molecular distances (R_{DA}) under unfolding conditions were estimated from measurements of time-resolved fluorescence resonance energy transfer between the donor Trp53 and the non-fluorescent acceptor TNB coupled to one of the seven cysteine side-chains. The unfolded protein chain expands with an increase in the concentration of the denaturants. The extent of expansion was found to be non-uniform, with different intra-molecular distances expanding to different extents. In general, shorter distances were found to expand less when compared to longer spans. The extent of expansion was higher in the case of GdnHCl when compared to urea. A comparison of the measured values of R_{DA} with those derived from a model based on excluded volume, revealed that while shorter spans showed good agreement, the experimental values of R_{DA} of longer spans were smaller when compared to the theoretical values. Sequence-specific flexibility of the polypeptide was determined by time-resolved fluorescence anisotropy decay measurements on acrylodan or 1,5-IAEDANS labeled single cysteine-containing proteins under unfolding conditions. Rotational dynamics derived from these measurements indicated that the level of flexibility increased with increase in the concentration of denaturants and showed a graded increase towards the C-terminal end. Taken together, these results appear to indicate the presence of specific non-random coil structures and show that the deviation from random coil structure is different for the two denaturants.

© 2006 Published by Elsevier Ltd.

Keywords: barstar; unfolded proteins; residual structure; protein dynamics; time-resolved fluorescence resonance energy transfer (TR-FRET)

*Corresponding authors

Abbreviations used: DTNB, 5,5'-dithio-bis(2-nitrobenzoic acid); TNB, thionitrobenzoate anion; TR-FRET, time-resolved fluorescence resonance energy transfer; acrylodan, 6-acryloyl-2-dimethylaminonaphthalene; 1,5-IAEDANS, 5-(((2-iodoacetyl)amino)ethyl)amino naphthalene-1-sulfonic acid; GdnHCl, guanidine hydrochloride; MEM, maximum entropy method; NATA, *N*-acetyl-1-tryptophanamide.

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

Introduction

In recent years, unfolded forms of proteins have attracted much attention for a variety of reasons, the foremost being the need to define better, the starting state of the protein folding reaction. The simplistic notion that proteins unfolded by chemical denaturants, such as urea and guanidine hydrochloride (GdnHCl), behave like random coils, has been challenged by the demonstration of residual structures in the unfolded forms of several proteins.^{1–5} Such observations, when coupled with the intuition that residual structures might arise from a sequence-specific propensity for native structure, raise the possibility of defining early events in protein folding from information obtained about unfolded proteins. For example, the nucleation–condensation mechanism for protein folding,^{6–8} is based implicitly, to some extent, upon such residual interactions which are hypothesized to persist in unfolded polypeptides. Thus, the seeds of the folding process may be sown in its unfolded form. Interest in unfolded forms of proteins also stems from the growing number of examples of proteins that appear to remain natively unfolded.^{9–11} The observation of disorder–order transitions even in such intrinsically unstructured proteins¹⁰ makes studies of the dynamics of unfolded proteins profitable. Finally, interest in unfolded proteins is augmented by observations that suggest that the unfolded form of a protein may play a seminal role in protein aggregation, including the formation of amyloid fibrils characteristic of protein misfolding disorders such as Alzheimer's, Parkinson's, and the prion diseases.^{12–14}

Structural characterization of unfolded proteins by high resolution nuclear magnetic resonance (NMR) spectroscopy¹⁵ has indicated, in several cases, the presence of hydrophobic clustering^{1–3} and fluctuating secondary structures.¹⁶ These structures could be stabilized by hydrophobic^{1,2,5} or electrostatic^{17–18} forces leading to either native-like¹⁹ or non-native-like²⁰ long range interactions. Residual secondary structure with some degree of compactness has been found in the thermally unfolded state of native ribonuclease A, by small-angle X-ray scattering (SAXS) and Fourier transform infrared spectroscopy.²¹ SAXS and circular dichroism (CD) measurements have shown an alteration in the conformational distribution of unfolded SNase delta by a single amino acid substitution.²² Statistical models when combined with experimental data provide a deeper understanding of the structural properties of polypeptide chains.²³ Although spectroscopic techniques such as NMR, IR, and CD have been used extensively to gather a wealth of information on the structures and dynamics of unfolded proteins, several limitations restrict their use. It is difficult to assign resonances in unfolded proteins by NMR, although there has been striking recent progress.²⁴ Furthermore, the low sensitivity of the NMR technique requires high protein concentrations, where many proteins tend

to aggregate in their unfolded form. IR and CD spectroscopy are not residue-specific, and hence cannot distinguish between various parts of the unfolded protein. On the other hand, measurements of the kinetics of intra-chain contact formation, by triplet–triplet energy transfer²⁵ and quenching of the triplet state of tryptophan by cysteine,^{26–27} have offered remarkable insights into the dynamics of unfolded proteins and peptides.

Time-resolved fluorescence spectroscopy is a highly sensitive and powerful tool to study residue-level dynamics,^{28–30} and to resolve structural heterogeneity in proteins.^{31–34} Time-resolved fluorescence resonance energy transfer (TR-FRET)^{32,35–36} measurements give information on the conformational heterogeneity in an ensemble of molecules through intra-molecular distance distributions. The efficiency of energy transfer between a donor–acceptor pair in a macromolecular system is highly sensitive to the distance between them. Hence, the distribution of distance between the donor and the acceptor chromophore will result in a distribution of energy transfer rates, which can be measured as a complex fluorescence intensity decay of the donor. Distributions of intra-molecular distances obtained in this way are similar to those obtained from single molecule FRET studies.^{37–39} Such a distance distribution-based approach has been used extensively to study conformational heterogeneity in protein folding,^{31–33} unfolding³⁶ and protein–protein association.³⁴ Another experimental tool for extracting site-specific information on the structure and dynamics of a fluctuating polypeptide chain is time-resolved fluorescence anisotropy. Various dynamic modes such as local, segmental and global motion can be visualized directly by time-resolved fluorescence anisotropy decay measurements of either tryptophan or a fluorophore tag.^{28–30} The degrees of freedom associated with the local and segmental dynamics in the unfolded chains can then be used to infer inter-residue interactions, segmental flexibility, and the presence of residual structure at that location.

There are only a few studies of unfolded proteins by site-specific time-resolved fluorescence spectroscopy. These studies have indicated the presence of residual structures,⁵ which are compact and possess high degrees of conformational heterogeneity.^{31,33,40} Due to their fluctuating nature and heterogeneous distributions, residual structures, which might serve as nucleation sites during the process of protein folding, are hard to identify by many other physical techniques.

Barstar, an 89 residue protein, which is the intracellular inhibitor of extracellular RNase barnase in *Bacillus amyloliquefaciens*, has been used extensively as a model system for protein folding studies.^{28,32,36,41} The unfolded form of barstar in high concentrations of urea appears to be devoid of native-like secondary structure,⁴² but electrostatic interactions appear to persist in it.⁴³ In order to understand the significance of the dynamics and structural features of unfolded proteins in general,

the present work focuses on characterizing chemically unfolded barstar. Seven single cysteine-containing mutants (Cys14, Cys25, Cys40, Cys42, Cys62, Cys82, Cys89) of barstar, each with a single tryptophan (Trp53) (Figure 1), have been used. Trp53 is located centrally in the hydrophobic core of the protein, positions 14, 25 are in helix-1, positions 40, 42 are in helix-2, position 62 is in helix-3, position 82 is in loop-6 and position 89 is in β -sheet-3 (see Figure 1).

Intra-molecular distance measurements by TR-FRET between Trp53 and various locations on chemically unfolded mutant variants of barstar show that the intra-molecular distances are sensitive to the concentration of the denaturant in the "unfolded protein-baseline" region. The results suggest an overall increase in the dimensions of the unfolded chain with an increase in the concentration of the denaturant. The extent of increase of an intra-molecular distance depends on the specific points on the polypeptide chain that are spanned by the distance, as well as on the nature of the denaturant. Residue-level rotational diffusion studies were also carried out by specifically labeling the mutant proteins at their single cysteine positions with a fluorescent probe, either 6-acryloyl-2-dimethylaminonaphthalene (acrylodan) or 5-(((2-iodoacetyl)amino) ethyl) amino)naphthalene-1-sulfonic acid (1,5-IAEDANS). Time-resolved fluorescence anisotropy measurements show that the level of rotational freedom is not uniform throughout the length of the polypeptide chain. Segments of the chain that are relatively less dynamic may have residual structures present and hence, higher

propensity for structure formation. This study attempts to relate the dynamics and structural features of the unfolded protein to its folding.

Results

Intra-molecular distances in unfolded barstar by TR-FRET

The estimation of intra-molecular distances in unfolded proteins by TR-FRET measurements is a sensitive way of probing conformational and structural heterogeneity.^{32,36} Intra-molecular distances between Trp53 (donor) and the non-fluorescent acceptor TNB covalently attached to the single cysteine in each of the seven single cysteine-containing mutant forms of barstar, in their denatured state, were determined by TR-FRET. Both unlabelled and TNB-labelled single cysteine-containing mutant proteins were unfolded by the denaturant urea (in the concentration range of 5 M to 8 M) or GdnHCl (in the concentration range of 3 M to 6 M). These concentration ranges fall within the flat unfolded protein-baseline regions of the respective denaturant-induced equilibrium unfolding curve of either cysteine-labelled or unlabelled barstar. The insets in Figure 2(a) and (b) show the urea and GdnHCl-induced equilibrium unfolding curves obtained for the unlabelled Cys82 protein, by monitoring the fluorescence of Trp53. All other single cysteine-containing mutant proteins, either unlabelled or labelled with TNB, showed similar unfolding curves, indicating their similar stability.

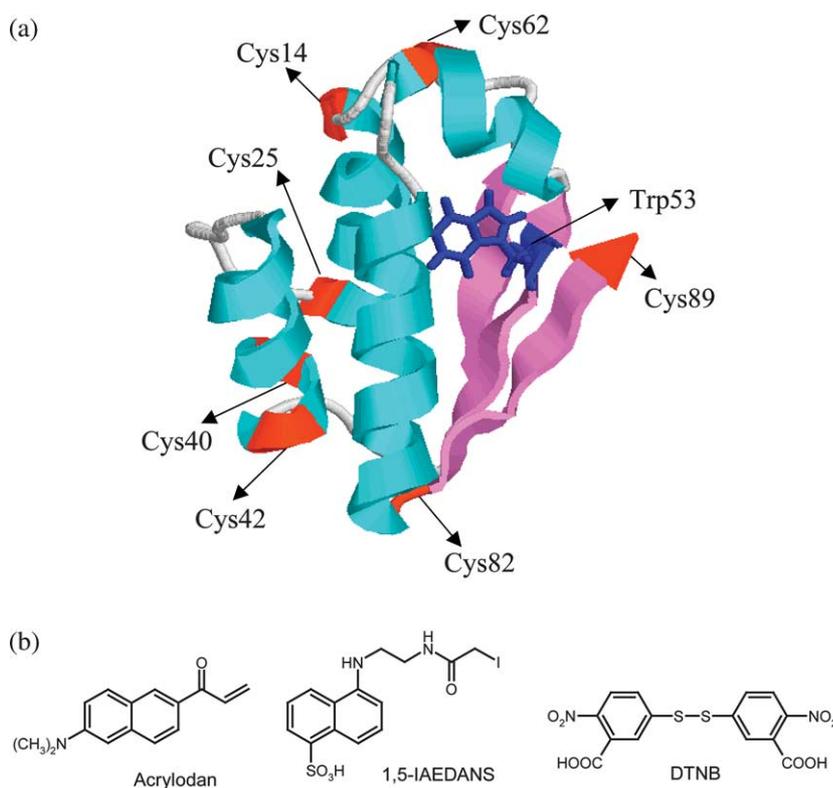


Figure 1. (a) Structure of barstar showing the locations of Trp53 and the various cysteine mutations. The structure was drawn using RAS-MOL software and the PDB file with accession code 1BTA. (b) The structures of various probes used to label cysteine residues are shown.

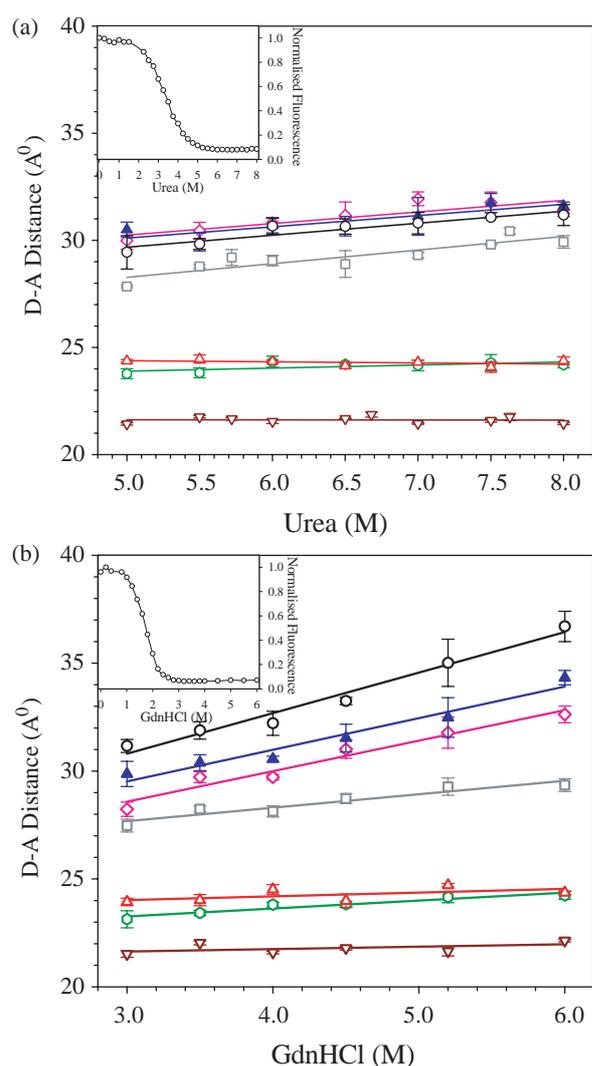


Figure 2. Denaturant concentration-dependence of distances (R_{DA}) between Trp53 and TNB labelled at various single cysteine sites in the seven single cysteine-containing mutant variants of barstar unfolded in (a) urea and (b) GdnHCl. Distances (Cys14TNB–Trp53), (Cys25TNB–Trp53), (Cys40TNB–Trp53), (Cys42TNB–Trp53), (Cys62TNB–Trp53), (Cys82TNB–Trp53) and (Cys89TNB–Trp53) are shown by open circles (black), open squares (dark gray), open hexagons (dark green), open up-triangles (red), open down-triangles (dark red), open diamonds (pink) and filled up-triangles (blue), respectively. The continuous line through the data points for each protein represents the linear regression fit. The error bars represent the standard deviations of measurements made in three separate experiments. The insets in (a) and (b) represent the equilibrium unfolding transitions of unlabelled single cysteine-containing variants of barstar, obtained by monitoring the fluorescence of Trp53 at 320 nm in urea and GdnHCl, respectively.

The mean fluorescence lifetime of Trp53, measured in the presence of the acceptor TNB covalently attached to a cysteine side-chain, was translated into intra-molecular distance, as described in Materials and Methods. Estimating

the distance from mean fluorescence lifetimes gives an average estimate of the intra-molecular distance, when the fluorescence decay has more than one component. The dependences of the distances between Trp53 and the seven single cysteine sites mentioned above, on the concentration of the denaturant in the range of concentrations corresponding to the unfolded baseline regions are shown in Figure 2. The slopes and correlation coefficients obtained from linear fits of these data are shown in Table 1. The following general features were observed for both the urea and GdnHCl-unfolded proteins. (i) The measured donor–acceptor distances (R_{DA}) increase with the increase in the concentration of denaturants, except for smaller separations which showed very little increase especially in the case of urea. (ii) The fractional increase in R_{DA} (slopes in Table 1) increases with increase in separations in the chain in most of the cases except for some distances (see later). (iii) The values of R_{DA} are similar at the highest concentration of urea and GdnHCl only for the shorter distances; and are different for the longer distances. (iv) The fractional increase in R_{DA} is higher in the case of GdnHCl-induced unfolding compared to urea-induced unfolding. Although the estimation of R_{DA} from changes in mean fluorescence lifetime (τ_m) is very similar to estimations based on steady-state fluorescence⁴¹ (since τ_m is proportional to fluorescence intensity), the former method is much more reliable, because errors in the determination of mean lifetimes are very small (Table 1). The use of τ_m instead of fluorescence intensity allows us to extend the range of estimation of R_{DA} to $\sim 1.6R_0$ with an acceptable level of accuracy (Table 1). The very high level of accuracy in the measurement of τ_m when coupled with its independence of factors such as concentration of the fluorophore, excitation intensity and collection efficiency allows us to estimate, reliably, even changes as low as 4% (Table 1). We used the randomized average value of 2/3 for the orientation factor, κ^2 since the levels of orientational freedom for the donor and acceptor are expected to be quite high in unfolded proteins. However, the range of κ^2 was estimated in typical cases by using the fluorescence anisotropy decay parameters associated with Trp53 and a fluorophore attached to acceptor sites.⁴⁴ For example, the range of values of R_{DA} estimated by taking into account the range of κ^2 (0.42–1.15) is 20.5–24.2 Å for Trp53–Cys62 in 6 M GdnHCl. Similar ranges were obtained for the other distances.

To determine whether the higher fractional increase in R_{DA} observed in the case of GdnHCl is due to the ionic character of the denaturant, the estimation of R_{DA} at various concentrations of KCl was carried out for a short intra-molecular distance (Cys40TNB–Trp53) and a longer distance (Cys82TNB–Trp53) in 8 M urea. Both the distances were found to be independent of the concentration of KCl up to 0.5 M (Figure 3). Thus, it appears that the greater fractional increase in R_{DA} that is observed with an increase in the concentration of

Table 1. Parameters obtained from the linear regression analysis of the denaturant dependence of various distances in the unfolded protein-baseline region of barstar

Segment on unfolded barstar	In GdnHCl					In urea		
	TR-FRET parameters (ns) (in 6 M GdnHCl) ^a		Distance (Å) (in 6 M GdnHCl) ^b	Slope of the regression line	Correlation coefficient ^c	Distance (Å) (in 8 M urea) ^b	Slope of the regression line	Correlation coefficient ^c
	τ_D	τ_{DA}						
Cys62TNB-Trp53	2.54±0.02	1.25±0.01	22.1±0.1	0.11	0.49	21.4±0.1	0.00	0.02
Cys42TNB-Trp53	2.54±0.02	1.61±0.01	24.4±0.1	0.17	0.60	24.4±0.2	0.00	0.43
Cys40TNB-Trp53	2.39±0.03	1.49±0.02	24.2±0.2	0.37	0.96	24.3±0.4	0.14	0.69
Cys25TNB-Trp53	2.54±0.01	2.13±0.02	29.4±0.3	0.62	0.95	29.9±0.4	0.64	0.89
Cys82TNB-Trp53	2.39±0.03	2.17±0.01	32.6±0.7	1.41	0.98	31.3±0.1	0.54	0.84
Cys89TNB-Trp53	2.50±0.01	2.33±0.01	34.3±0.4	1.46	0.98	31.6±0.2	0.53	0.89
Cys14TNB-Trp53	2.50±0.01	2.38±0.01	36.7±0.7	1.90	0.99	31.2±0.5	0.56	0.94

^a τ_D and τ_{DA} are the mean fluorescence lifetimes of Trp53 in the absence and presence of acceptor TNB in barstar unfolded in 6 M GdnHCl. The errors in the values of the lifetimes represent the standard deviations of measurements made in three separate experiments.

^b Distances between Trp53 and Cys-TNB in barstar, measured by TR-FRET.

^c Correlation coefficient of the linear regression analysis.

GdnHCl, as compared to that observed with an increase in concentration of urea (Figure 2), reflects the differential effects of GdnHCl and urea as denaturants, rather than the ionic nature of GdnHCl.

When the concentration of denaturant is increased from 5 M to 8 M urea, or from 3 M to 6 M GdnHCl, the solvent viscosity increases by factors of 1.14 and 1.21, respectively. An increase in solvent viscosity will slow down Brownian diffusion of the unfolded polypeptide chain, and hence,

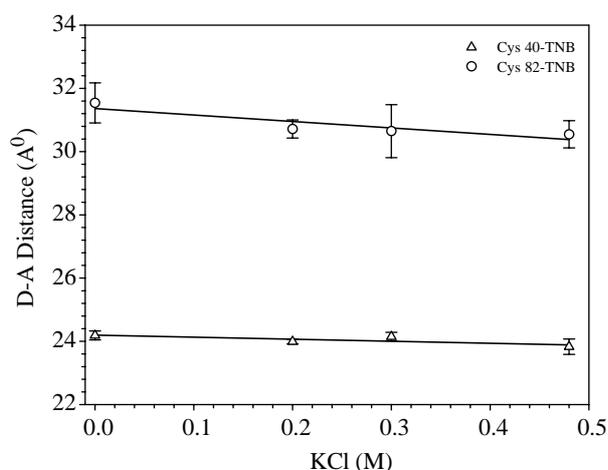


Figure 3. Dependence on the concentration of KCl, of the distances between Trp53 and TNB-labelled Cys40 (shown by open triangles) and TNB-labelled Cys82 (shown by open circles) proteins unfolded in 8 M urea. The continuous line through the data points for each protein represents the linear regression fit.

might affect TR-FRET-based distance measurements. During the excited state lifetime of the donor, Brownian diffusion would cause fluctuation in R_{DA} , especially for flexible systems such as unfolded proteins. Due to the inverse sixth power dependence of energy transfer rate on R_{DA} , the estimates of distribution of R_{DA} would be biased towards the lower limit of the distribution of R_{DA} generated by the Brownian diffusion-driven changes in the value of R_{DA} during the excited state lifetime of the donor.⁴⁵ Based on this model, there have been attempts to estimate the intra-chain translational diffusion coefficient in flexible polypeptides.⁴⁵ Since an increase in solvent viscosity is expected to slow down Brownian diffusion, and hence, reduce the width of such distributions of R_{DA} , the apparent value of R_{DA} could be expected to increase with the increase in solvent viscosity. To check whether the observed increase in R_{DA} with the increase in the concentration of either urea or GdnHCl is caused by the increase in viscosity, estimation of R_{DA} was carried out at both 5 M urea and 3 M GdnHCl with appropriate amounts of glycerol added such that the solvent viscosity matched those of 8 M urea and 6 M GdnHCl, respectively. It was found that the values of both τ_{DA} and τ_D increased almost by the same factor (data not shown), upon changing the viscosity of the 5 M urea solution to that of an 8 M urea solution or of a 3 M GdnHCl solution to that of a 6 M GdnHCl solution by adding glycerol. This result indicates that these changes are due merely to an increase in the quantum yield of Trp53 caused by the addition of glycerol. Since the value of R_0 remains essentially unchanged, it appears that the

value of R_{DA} does not change appreciably upon an increase in viscosity up to ~ 1.2 cP.

Distributions of energy transfer kinetics in unfolded barstar

The distribution of fluorescence lifetime of Trp53 in the presence of acceptor TNB, τ_{DA} , in mutant proteins of barstar were determined by the maximum entropy method (MEM) of analysis of fluorescence intensity decay kinetics,^{36,46–48} as described in Materials and Methods. Figure 4 shows typical MEM distributions of τ_{DA} of the TNB-labelled proteins and τ_D of an unlabelled protein all unfolded in 8 M urea. Similar distributions were observed in the case of proteins denatured by GdnHCl (data not shown). Multiple peaks observed in these distributions (in both labelled and unlabelled proteins) are likely to be due to the conformational heterogeneity arising from the distribution of population of rotamers around the $C^\alpha-C^\beta$ bond in the tryptophan side-chain.⁴⁷ The heterogeneity could also be due to

interaction of adjacent side-chains with the indole ring of the tryptophan. The distributions corresponding to the native states of the proteins are also shown (Figure 4). These distributions are nearly unimodal, unlike those of the unfolded proteins, and are due to conformational restriction of Trp53 in the N-state, and they represent a unique situation among several single-tryptophan proteins, including barstar.^{28,36} For some of the unfolded proteins, the distributions of τ_{DA} display minor peaks in the short lifetime (~ 0.1 ns) region where the distributions of τ_{DA} for the folded proteins display (Figure 4) their unimodal peaks.

Rotational dynamics of fluorescence probes in unfolded barstar

In order to further address the question whether proteins unfolded by denaturants show true local random coil structural characteristics, the degree of flexibility at various locations on the unfolded form of barstar was studied by measurements of the

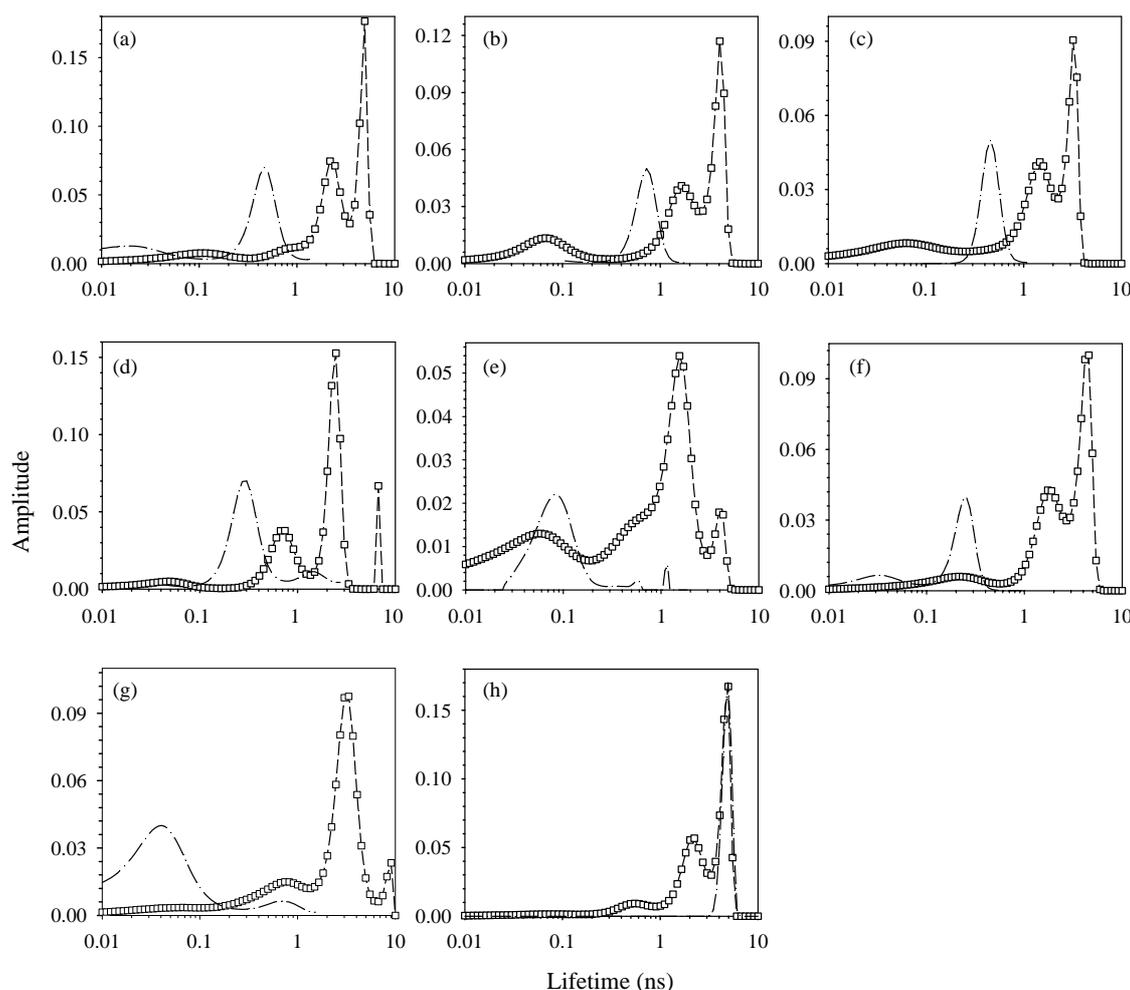


Figure 4. Fluorescence lifetime distributions of Trp53 in (a) Cys14-TNB, (b) Cys25-TNB, (c) Cys40-TNB, (d) Cys42-TNB, (e) Cys62-TNB, (f) Cys82-TNB, (g) Cys89-TNB, and (h) unlabelled Cys82 proteins obtained by the maximum entropy method (MEM) of analysis. Dash-dot-dash (– • –) line represents the protein in the native form and the broken line with open square symbols (– □ –) represents the protein unfolded in 8 M urea.

rotational dynamics of fluorophores attached to the single cysteine side-chain of the mutant proteins. Thiol-specific fluorophores, acrylodan and 1,5-IAEDANS were used as molecular probes. The labelled proteins were unfolded by urea (at concentrations of 5 M and 8 M) and GdnHCl (at concentrations of 3 M and 6 M) and the decay kinetics of fluorescence anisotropy were determined.

Typical decays of fluorescence anisotropy of acrylodan-labelled mutant proteins in 8 M urea and 6 M GdnHCl are shown in Figure 5. The decay curves observed at various locations are found to be dissimilar in terms of their kinetics. For all the proteins, the decay kinetics could be fitted to a model with two correlation times (equation (5)), where the fast correlation time (ϕ_2) represents the local motional freedom of the probe with respect to the polypeptide chain and the slower correlation time (ϕ_1) represents a combination of the segmental mobility of the polypeptide region with respect to

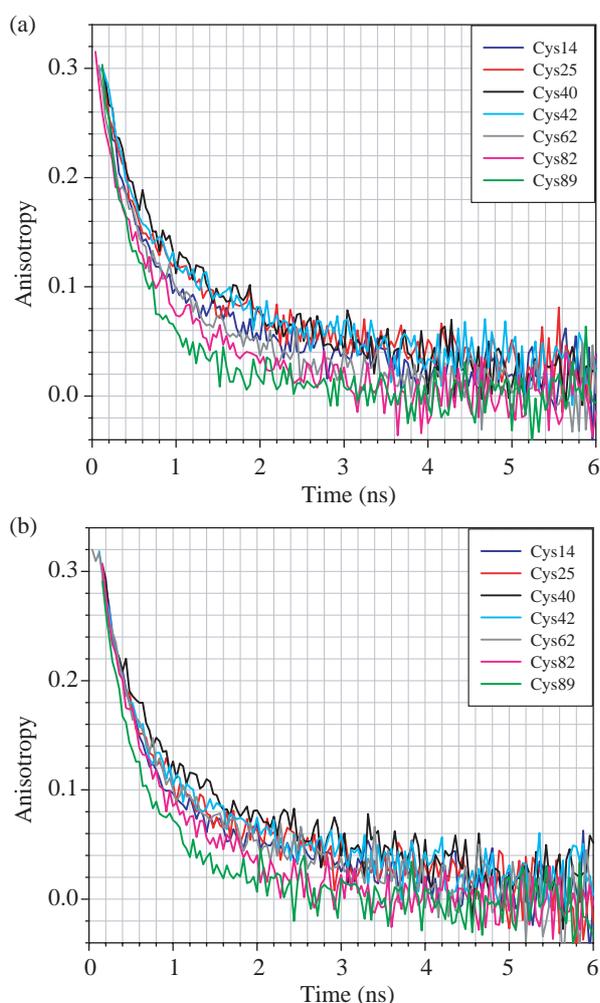


Figure 5. Fluorescence anisotropy decay profiles of acrylodan-labelled single-cysteine variants of barstar unfolded in (a) 8 M urea and (b) 6 M GdnHCl. The emission was monitored at 510 nm. Rotational correlation times estimated from these decay profiles are given in Table 2.

the overall chain and the global tumbling dynamics of the entire chain. The values of the rotational correlation times (ϕ_1 and ϕ_2) obtained from analysis of the decay kinetics of acrylodan-labelled mutants of barstar under various unfolding conditions are listed in Table 2. Table 3 gives the parameters obtained from mutant proteins labelled with 1,5-IAEDANS.

It is seen from Tables 2 and 3 that almost all parameters (correlations times and their amplitudes) show a dependence on the position of the fluorophore along the sequence of the unfolded protein. The following general observations are relevant. (i) The overall flexibility shows an increase with an increase in the concentration of the denaturant in the unfolded protein baseline region (5–8 M of urea or 3–6 M GdnHCl). This conclusion comes mainly from the observed decrease, at all the positions of acrylodan-labelled proteins, in the value of ϕ_1 with the increase in the concentration of denaturants. In the case of the mutant proteins labelled with 1,5-IAEDANS, the apparent value of ϕ_1 does not show any decrease in contrast to the observation in acrylodan-labelled proteins. It is, however, to be noted that the solvent viscosity increases by 1.14 and 1.21-fold when the concentrations of urea and GdnHCl are increased from 5 M to 8 M and from 3 M to 6 M, respectively. Since ϕ_1 is likely to have contributions from global dynamics, its value could scale with viscosity as demanded by the Stokes–Einstein hydrodynamics equation. Furthermore, segmental dynamics might also depend upon the solvent viscosity. Thus, the actual value of ϕ_1 at each of the higher denaturant concentrations should be lower (by the factors mentioned) than the apparent values listed in Tables 2 and 3. These changes would then result in a general observation that, for both the probes, the value of ϕ_1 decreases with an increase in the concentration of denaturant. (ii) The level of flexibility of the C-terminal region (82–89) is significantly higher when compared to the rest of the chain as shown by the value of ϕ_1 for the acrylodan-labelled proteins. (iii) The shorter correlation time, ϕ_2 which represents the local dynamics of the probe, and its amplitude, does not show any appreciable dependence on either the concentration of the denaturants, or the position of the probe. The dependence of local (internal) dynamics on solvent viscosity is more complex and case-dependent⁴⁹ and hence, its correction for solvent viscosity is not attempted. (iv) The extent of denaturant-induced decrease in the value of ϕ_1 is larger in the case of acrylodan when compared to 1,5-IAEDANS. This could be due to the shorter separation of the chromophore from the peptide backbone, in the case of acrylodan enabling it to sense better, the level of structure formation by the peptide backbone. This assertion is supported by the observation that the amplitude (β_2) associated with the local motion (shorter correlation time, ϕ_2) is smaller for acrylodan when compared to 1,5-IAEDANS for all the positions except Cys14, indicating the increased

Table 2. Parameters associated with fluorescence anisotropy decay in acrylodan-labelled single cysteine-containing mutants of barstar, obtained under various denaturing conditions at room temperature

Barstar mutant forms labelled at cysteine	Rotational correlation times (ϕ) and amplitudes (β)							
	5 M Urea		8 M Urea		3 M GdnHCl		6 M GdnHCl	
	$\phi_1 \pm 0.5$ (β_1)	$\phi_2 \pm 0.05$ (β_2)	$\phi_1 \pm 0.3$ (β_1)	$\phi_2 \pm 0.05$ (β_2)	$\phi_1 \pm 0.5$ (β_1)	$\phi_2 \pm 0.05$ (β_2)	$\phi_1 \pm 0.3$ (β_1)	$\phi_2 \pm 0.05$ (β_2)
Cys14-acrylodan	4.00 (0.39)	0.32 (0.61)	2.30 (0.40)	0.32 (0.60)	3.35 (0.38)	0.24 (0.62)	1.90 (0.38)	0.23 (0.62)
Cys25-acrylodan	4.17 (0.51)	0.33 (0.49)	3.00 (0.41)	0.41 (0.59)	4.10 (0.44)	0.33 (0.56)	2.12 (0.43)	0.34 (0.57)
Cys40-acrylodan	3.70 (0.54)	0.35 (0.46)	2.68 (0.47)	0.38 (0.53)	3.80 (0.53)	0.31 (0.47)	2.36 (0.54)	0.28 (0.46)
Cys42-acrylodan	4.58 (0.46)	0.32 (0.54)	2.77 (0.45)	0.33 (0.55)	4.07 (0.48)	0.28 (0.52)	2.47 (0.41)	0.34 (0.59)
Cys62-acrylodan	2.38 (0.44)	0.31 (0.56)	1.67 (0.51)	0.30 (0.49)	3.16 (0.41)	0.36 (0.59)	1.61 (0.53)	0.35 (0.47)
Cys82-acrylodan	2.57 (0.34)	0.35 (0.66)	1.05 (0.61)	0.25 (0.39)	3.70 (0.54)	0.33 (0.46)	1.15 (0.52)	0.36 (0.48)
Cys89-acrylodan	2.20 (0.23)	0.30 (0.77)	1.09 (0.34)	0.32 (0.66)	3.60 (0.23)	0.32 (0.77)	1.33 (0.26)	0.37 (0.74)

The error in β_1 and β_2 is 0.02.

level of restriction in the case of acrylodan (Tables 2 and 3).

Discussion

The dimensions of unfolded barstar depend on the concentration of denaturants

Ever since Tanford⁵⁰ showed that several proteins exhibit the hydrodynamic properties of random coils in the presence of denaturants, the question as to whether the unfolded state is a true random coil

or not, under all conditions, has been debated strongly.^{23,51–55} Interest in this problem stems mainly from the recognition that any deviation from random coil behavior can have implications regarding the initial events of protein folding, and may provide a rational hypothesis to construct the pathway of folding. Implied in this expectation is that any deviation from random coil behavior could be sequence-specific, and might enable the prediction of the pathway of folding for a given sequence.

Of the several physical techniques being used to map the conformational dynamics of unfolded proteins,⁵⁶ small angle X-ray and neutron scattering

Table 3. Parameters associated with fluorescence anisotropy decay in 1,5-IAEDANS-labelled single cysteine-containing variants of barstar, which were obtained under various denaturing conditions at room temperature

Barstar mutant forms labelled at cysteine	Rotational correlation times (ϕ) and amplitudes (β)							
	5 M Urea		8 M Urea		3 M GdnHCl		6 M GdnHCl	
	$\phi_1 \pm 0.5$ (β_1)	$\phi_2 \pm 0.05$ (β_2)	$\phi_1 \pm 0.5$ (β_1)	$\phi_2 \pm 0.05$ (β_2)	$\phi_1 \pm 0.5$ (β_1)	$\phi_2 \pm 0.05$ (β_2)	$\phi_1 \pm 0.5$ (β_1)	$\phi_2 \pm 0.05$ (β_2)
Cys14-IAE-DANS	4.10 (0.40)	0.25 (0.60)	2.65 (0.52)	0.37 (0.48)	2.42 (0.45)	0.34 (0.55)	2.50 (0.46)	0.40 (0.54)
Cys25-IAE-DANS	2.90 (0.43)	0.42 (0.57)	2.70 (0.44)	0.45 (0.56)	2.85 (0.37)	0.55 (0.63)	3.00 (0.35)	0.60 (0.65)
Cys40-IAE-DANS	3.00 (0.37)	0.45 (0.63)	3.00 (0.38)	0.46 (0.62)	2.62 (0.46)	0.36 (0.54)	2.80 (0.44)	0.40 (0.56)
Cys42-IAE-DANS	3.00 (0.37)	0.35 (0.63)	3.00 (0.37)	0.44 (0.63)	3.00 (0.40)	0.37 (0.60)	3.00 (0.41)	0.48 (0.59)
Cys62-IAE-DANS	2.10 (0.42)	0.35 (0.58)	2.50 (0.40)	0.42 (0.60)	2.45 (0.36)	0.40 (0.64)	2.40 (0.45)	0.40 (0.55)
Cys82-IAE-DANS	2.10 (0.30)	0.35 (0.70)	2.02 (0.33)	0.43 (0.67)	2.05 (0.30)	0.40 (0.70)	2.05 (0.35)	0.42 (0.65)
Cys89-IAE-DANS	3.00 (0.15)	0.35 (0.85)	3.00 (0.17)	0.38 (0.83)	3.00 (0.18)	0.35 (0.82)	3.00 (0.20)	0.35 (0.80)

The error in β_1 and β_2 is 0.02.

(SAXS and SANS, respectively) techniques and FRET-based methods are very powerful in providing the overall dimension through radius of gyration (R_G) and point-to-point (R_{DA}) distances, respectively. Although there has not been any direct comparison of the solvent dependence of R_G and R_{DA} for any protein, it is nevertheless useful to point out that these two indicators of molecular dimension provide complementary information. R_G is a measure of the overall dimension of the molecule and could be insensitive to the presence of intra-molecular interactions leading to stiffness.⁵⁷ In contrast, R_{DA} is more specific, and hence, it is likely to be modulated by the level of inter-residue contacts in the relevant span of the unfolded protein.

A principal result of this work, in which seven intra-molecular distances (R_{DA}), in barstar unfolded by either urea or GdnHCl, were estimated by time-resolved FRET, is that these distances expand with an increase in the concentration of either denaturant, that different distances expand differentially, and that the degree of expansion is dependent on the nature of the denaturant. The larger values observed for changes in R_{DA} in the cases of longer sequence segments (such as Trp53–Cys89) compared to shorter sequence segments (such as Trp53–Cys40; Figure 2) might indicate a uniform expansion of the unfolded protein, but different long sequence segments are seen not to expand similarly.

The expansion may be the result of domination of solvent–peptide interactions over inter-residue interactions when the concentration of denaturants is increased. In fact, statistical mechanics theory of protein stability predicts that the dimensions of weakly hydrophobic sequences will depend upon the nature of the solvent, and will increase as the denaturing power of the solvent increases.^{58–59} These studies also predict that the changes in the unfolded state will be non-cooperative, similar to the observations reported here (Figure 2). The unfolded form comprises molecules with differing dimensions, and it is likely that increased binding of the denaturant to the polypeptide chain, with an increase in the concentration of the denaturant, leads to preferential stabilization of the more expanded molecules. In this way, the transition from less expanded to more expanded forms, is expected to be a continuous transition. When the weak binding of denaturant to the polypeptide chain⁶⁰ saturates at high denaturant concentration, further expansion will stop, and the dimensions of the chain are expected to reach a plateau level. This appears to be the case for several of the intra-molecular distances being individually monitored in the present study. It should be noted, however, that the observation that different long spans expand differently with an increase in denaturant concentration (Figure 2 and Table 1), suggests that observed effects of denaturant, particularly GdnHCl, on the average separation, are not consistent with a model based simply on improved solvation of the polypeptide chain.

Single molecule FRET experiments of unfolded cold shock protein CspTm,³⁸ chymotrypsin inhibitor 2⁶¹ and ribonuclease H⁶² have shown that these unfolded proteins too expand with an increase in GdnHCl concentration. In the case of CspTm,³⁸ expansion appears to level off at high concentrations of denaturant, a result also seen in SAXS measurements of other proteins at high concentrations of denaturant, showing that R_G is largely insensitive to both the type and concentration of chemical denaturants.^{55,63–67} Ensemble FRET studies of unfolded cytochrome *c*⁶⁸ also indicate that the unfolded protein expands with an increase in the concentration of GdnHCl. Similarly, SAXS studies of the B1 domain of streptococcal protein G show also that R_G of the unfolded form of the *m*⁻ mutant protein increases with the concentration of the denaturant.⁶⁹ Estimates of the rates of intra-molecular contact formation in model peptides²⁵ as well as the unfolded forms of CspTm²⁶ and the headpiece sub-domain of villin²⁷ indicate that the rates increase with a decrease in the concentration of denaturants, suggesting an increased level of compaction at lower concentrations of denaturants. The probability of loop formation in unfolded *iso*-1-cytochrome *c* has also been shown to depend on the concentration of denaturant.⁷⁰ Thus, it appears that denaturant-induced expansion of unfolded proteins could be a general property of unfolded polypeptide chains, and a mean field theory model, in which solvophobic interactions are pitched against chain entropy, predicts just this.⁵⁹ Nevertheless, for both ubiquitin and acylphosphatase, the dimensions of the unfolded form do not appear to change with a change in denaturant concentration, even at low concentrations of the denaturant.⁷¹

GdnHCl and urea-unfolded barstar behave differently

The question whether proteins denatured by either GdnHCl or urea are similar in their structure and dynamics has not been resolved, although a protein is generally believed to be a random coil in either solvent. The present study is one of a few studies where the effects of urea and GdnHCl on the unfolded state of a protein have been compared directly. A higher level of expansion is observed in the case of GdnHCl when compared to the changes in urea, and the reason why this cannot be due to the salt effect of GdnHCl has been discussed already. The difference between the behaviors of the two denaturants is not surprising because the nature of their interaction with the polypeptide need not be the same. A calorimetric study of the interaction of GdnHCl and urea with several proteins had, in fact, shown that the number of binding sites of urea is two to threefold higher when compared to that of GdnHCl in both the native and denatured states.⁷² Based on this, it was proposed that GdnHCl forms nearly a twofold higher number of H-bonds with protein groups when compared to

urea.⁷² Furthermore, binding constants of GdnHCl and urea to peptides estimated from denaturant-dependence of the end-to-end diffusion of model peptides are 0.62 M^{-1} and 0.26 M^{-1} for GdnHCl and urea, respectively.⁶⁰ This difference and the difference in the mode of H-bond formation might be the origin of the differential effects observed in this work. Characterization of the structure and dynamics of unfolded FK506-binding protein by NMR had revealed subtle differences regarding the location of persistent secondary structures in the protein unfolded by either GdnHCl or urea.⁴ The differences in barstar unfolded in urea and GdnHCl, might be responsible for why the folding and unfolding pathways of the protein appear to be different in the two different denaturants.⁷³

Is barstar a true random coil in high concentrations of denaturant?

The dependence of the dimensions of a polymer on its length is a key criterion for the diagnosis of the randomness of polymer structure. Early work by Flory⁷⁴ predicted that, under favourable solvation conditions, the RMS values of R_{DA} and R_G will be proportional to $N^{3/5}$ for a chain of N monomers when excluded volume is taken into account. Measurements of R_G by SAXS, for the unfolded forms of many proteins with varying length, follow the random coil behavior with an exponent of 0.6.⁶³ Similarly, Monte Carlo simulations of several polypeptides whose conformations were restricted only by steric repulsions, showed that both the end-to-end distance and R_G scale with an exponent of 0.583,⁷⁵ which is close to the value obtained by Flory. Thus, a comparison between the experimentally observed values of the end-to-end distance and the radius of gyration, and values predicted by random coil simulations⁷⁵ can be used as a criterion for identifying any deviation from a true random coil. It should, however, be pointed out that observation of random coil behavior does not exclude the presence of stiff segments in the chain.⁵⁷

Table 4 presents the values of R_{DA} obtained experimentally in 6 M GdnHCl (Figure 2). The values calculated from the equation $R_{coil} = 5.68n^{0.583}$

(where $n = a_1 - a_2$; a_1 and a_2 are the positions of the respective amino acids) are given for comparison. As mentioned earlier, this equation was deduced from Monte Carlo simulations⁷⁵ on various polypeptides whose conformations are restricted only by steric repulsion, and hence, would represent true random coils. It can be seen that the level of agreement between the measured values of R_{DA} and the calculated random coil values R_{coil} is rather mixed. There is a fairly good level of agreement in the cases of shorter distances (Cys40–Trp53), (Cys42–Trp53), and (Cys62–Trp53), over the entire range of denaturant concentrations studied (Figure 2). On the other hand, the longer spans are highly underestimated by experiments when compared to the random coil values; moreover, the degree of underestimation is greater at the lower concentrations of denaturant, because these distances are seen to decrease with decreasing denaturant concentration (Figure 2). The deviation from the expected random coil distances of the distances corresponding to the longer spans, especially at lower concentrations of denaturant where the protein is still fully unfolded, is indicative of non-random coil structure.

The calculated values of R_{coil} represent the distance between the sulphur atoms of the pair of cysteine residues.⁷⁵ To compare the values of R_{coil} with the distances (R_{DA}) measured by FRET experiments, values of R_{coil} should be corrected for the dimensions of the donor and acceptor probes used in FRET experiments. The flexibility of the probes would result in uncertainty in the values of R_{coil} . The maximum value of R_{coil} would correspond to the conformation in which the emission dipole of the donor and the absorption dipole of the acceptor probes are maximally elongated along the line connecting the donor and acceptor. The minimum value would correspond to the conformation where the dipoles are oriented towards each other. The distance between the γ -carbon and the centre of the indole side-chain of the tryptophan residue (donor) is $\sim 1.8 \text{ \AA}$ and the distance between the sulphur atom at the γ -position of the cysteine residue and the centre of the benzene ring of the TNB (acceptor) labelled to the cysteine thiol is

Table 4. Calculated and measured values of distances between Trp53 and various single cysteine sites on unfolded mutant proteins of barstar

Segment on unfolded barstar	n^a	$R_{coil} (\text{\AA})^b$	$R_{coil(\min)} (\text{\AA})^c$	$R_{coil(\max)} (\text{\AA})^d$	$R_{DA(\max)} (\text{\AA})^e$
Cys14TNB–Trp53	39	48.1	44.5	54.3	36.7 ± 0.7
Cys25TNB–Trp53	28	39.6	36.0	45.8	29.4 ± 0.3
Cys40TNB–Trp53	13	25.3	21.7	31.5	24.2 ± 0.2
Cys42TNB–Trp53	11	23.0	19.4	29.2	24.4 ± 0.1
Cys62TNB–Trp53	09	20.5	16.9	26.7	22.1 ± 0.1
Cys82TNB–Trp53	29	40.5	36.9	46.7	32.6 ± 0.7
Cys89TNB–Trp53	36	45.9	42.3	52.1	34.3 ± 0.4

^a $n = (a_1 - a_2)$, where a_1 and a_2 are the positions of the respective amino acids.

^b Random coil distance $R_{coil} = (5.68 \pm 0.08)n^{(0.583)} \text{ \AA}$.⁷⁵

^c Minimum random coil distance $R_{coil(\min)} \approx (R_{coil} - 3.6) \text{ \AA}$, see the text.

^d Maximum random coil distance $R_{coil(\max)} \approx (R_{coil} + 6.2) \text{ \AA}$, see the text.

^e $R_{DA(\max)}$ is the maximum distance between the FRET pair in barstar unfolded in 6 M GdnHCl, obtained by TR-FRET experiments.

~ 4.4 Å. Hence, the maximum extent of correction to R_{coil} would be 6.2 Å. When the dipoles point towards each other, their projections along the axis of the polypeptide segment are estimated to be ~ 1.1 Å and ~ 2.5 Å respectively, resulting in a negative correction of 3.6 Å to R_{coil} . A comparison of the estimated values of R_{DA} and the range of random coil values R_{coil} , with these considerations taken into account, leads, once again, to the conclusion mentioned above, viz. there is agreement in the short spans but not in the longer spans.

The disagreement between the experimental and calculated values of longer distances in unfolded barstar might be due to the following reasons. (i) A FRET efficiency higher than expected from Forster's theory with random orientations of D and A could be due to insufficient averaging of orientations during the lifetime of the donor, and also due to breakdown of point dipole approximation used;³⁷ however, these effects are expected to result in disagreement in all the distances contrary to our observations. (ii) Segmental Brownian diffusion of the polypeptide chain during the donor lifetime could have caused the reduction in the values of R_{DA} .⁴⁵ This effect would be more dominant in the distance estimation of longer spans of the polypeptide chain due to a larger degree of translational freedom. However, this effect is likely to be insignificant due to the short lifetime of the tryptophan donor.³¹ (iii) Non-random structure could be present in the regions spanned by the longer distances. The observation that several intramolecular distances, especially those associated with larger spans, show non-saturating behavior in their increase (Figure 2) indicates that they are yet to reach the random coil values. This alone could be taken as a strong support for the assertion that the structure of unfolded barstar in the presence of denaturants shows specific deviations from a random coil.

In the case of the cold shock protein *Bc-Csp*, good overall agreement between several FRET-determined values of R_{DA} in unfolded protein and the random coil values, R_{coil} was reported, and it was argued that the unfolded protein is devoid of any residual structures.⁷⁶ In that study, random coil distances were, however, determined using $R_{\text{coil}} = 5.45n^{0.5}$ Å, which is applicable to random chains with the neglect of excluded volume. A more appropriate value for the exponent is 0.6, which results from the analytical theory of Flory⁷⁴ or 0.583, which has been determined from Monte Carlo simulations on random coil polypeptides,⁷⁵ in which excluded volume effects are incorporated. We note that in our study also, several experimental values of R_{DA} (such as Trp53–Cys25 and Trp53–Cys89) show a better match with R_{coil} estimated by $R_{\text{coil}} = 5.45n^{0.5}$ Å, but the validity of the model is obviously questionable. In the case of the unfolded FynSH3 domain, site-specific deviations of the FRET-determined values of R_{DA} from the random coil distances, R_{coil} , have also been observed.³⁹

Features of non-random structure in unfolded barstar

Characterization of the residue-specific rotational dynamics of covalently linked fluorophores (Figure 5 and Tables 2 and 3) shows that the level of segmental and global dynamics increases with an increase in the concentration of denaturants at all the seven locations monitored in this study. An increase in the level of rotational dynamics of probes is generally a result of increased flexibility and hence, a decreased level of inter-residue interactions. Thus, it appears that inter-residue interactions are prevalent throughout the sequence of unfolded barstar.

In addition to such a general picture, the present study also provides information on some specific aspects of residual structures. The level of flexibility shows a graded increase towards the C-terminal region as revealed by the parameters associated with the rotational dynamics, especially the rotational correlation time, ϕ_1 which is ascribed to represent predominantly the segmental dynamics of polypeptides (Tables 2 and 3). This graded increase in flexibility is seen with both the fluorescence probes (1,5-IAEDANS and acrylodan), and with both the denaturants, and is therefore a very robust observation. Since inter-residue interactions could be the main cause of reduced flexibility of a segment of polypeptide, this result appears to suggest a higher prevalence of inter-residue interactions, and hence, residual structures in regions closer to the N-terminal region. Structurally constrained regions such as loops may play an important role in early events of folding.⁷⁰ The stiffness and length of such structurally constrained regions in an unfolded polypeptide chain would decide the rate of intra-chain diffusion necessary for the loop formation⁷⁷ which sets an upper limit to the folding rate.^{25,78}

It should be noted, however, that an NMR study of the urea-unfolded form of barstar has shown it to be devoid of native-like secondary structure, but has shown that residues 83 to 89 at the C-terminal region, corresponding to the last β -strand in the native protein, may adopt non-native helical conformations in 8 M urea.⁴² On the other hand, another NMR study, of cold-denatured barstar, has indicated that the cold-denatured protein does not possess structure in the C-terminal region Asn65 to Ser89.⁷⁹ The NMR experiments therefore also suggest, like the results reported here, that the C-terminal region in barstar is flexible.

By the determination of specific intra-molecular distances, as well as by the determination of rotational dynamics at specific sites on the polypeptide chain, it has been shown here that structure and dynamics vary along the length of the chain. The correlation between structure and dynamics in the unfolded protein is, however, not easy to determine, and more such studies on other proteins are needed to further understand these important aspects of the unfolded form of a protein.

Are compact structures present in the ensemble of unfolded protein molecules?

The demonstration of the presence of residual structures in unfolded barstar raises the question whether any component in the population of unfolded protein molecules could be a compact structure with dimensions similar to that of the native protein. Winkler and co-workers⁴⁰ estimated distributions of various intra-molecular distances in unfolded yeast *iso-1* cytochrome *c* by MEM analysis of the kinetics of fluorescence energy-transfer from a dansyl fluorophore to the Fe(III) heme. The distributions showed multiple peaks, with some distances very close to those of native protein. These were interpreted as compact structures in the unfolded ensemble. In that study, however, distributions of R_{DA} were constructed by assuming a single value of lifetime of the donor in the absence of acceptor as obtained from dansyl-modified *N*-acetylcysteine, rather than from a donor attached to the protein.^{40,80–81} This approximation is unlikely to be valid due to the high sensitivity of fluorescence lifetimes to environment, and could have caused the artifactual recovery of native-like distances for unfolded yeast *iso-1* cytochrome *c*. In this TR-FRET study of unfolded barstar, the donor Trp53 was part of the polypeptide chain, and the measurement of fluorescence lifetime distributions (Figure 4) in seven mutant proteins indicates that the fraction of unfolded molecules possessing native-like compactness is insignificant in the total population of unfolded molecules.

Materials and Methods

Chemicals

All the chemicals used were of the highest purity grade available from Sigma Aldrich Inc. 1,5-IAEDANS and acrylodan were from Molecular Probes. The buffer used in all the experiments was 20 mM Tris (pH 8.0), 250 μ M EDTA. Concentrations of urea and GdnHCl stock solutions were determined by refractive index measurements. The protein concentration used was 5 μ M. All the measurements were done at 25 °C.

Protein purification and labelling

Wild-type barstar contains three tryptophan residues (Trp38, Trp44 and Trp53) and two cysteine residues (Cys40 and Cys82). All the mutant variants of barstar used here were generated by site-directed mutagenesis. Each mutant protein contains a single tryptophan (Trp53) residue and a single cysteine residue located at one of seven different locations on the protein. For simplicity, the mutant proteins W38F/W44F/C40A/C82A/S14C, W38F/W44F/C40A/C82A/A25C, W38F/W44F/C82A, W38F/W44F/C40A/C82A/T42C, W38F/W44F/C40A/C82A/L62C, W38F/W44F/S12T/C40A, and W38F/W44F/C40A/C82A/S89C, are denoted by the position of the single cysteine residue present in them as Cys14, Cys25, Cys40, Cys42, Cys62, Cys82, and Cys89, respectively. All these proteins were purified as described⁸² and

their purity was confirmed to be >98% on SDS-PAGE. The protein concentrations were determined by absorbance at 280 nm, using a value for ϵ_{280} of 10,000 $M^{-1} cm^{-1}$ for all the proteins.⁸³

For TR-FRET measurements, all the mutant proteins were labelled at their single cysteine site with TNB by reacting with a 100-fold molar excess of DTNB in 8 M urea at pH 8.5. After the labelling reaction was complete, the labelled protein was separated from free dye and urea by passing the reaction mixture through a PD 10 column. The extent of labelling was confirmed to be >98% by a DTT assay.⁸⁴ Similarly, for time-resolved anisotropy measurements, all the mutant proteins were labelled with 1,5-IAEDANS or acrylodan in 8 M urea at pH 8 with about a 10 to 15-fold molar excess of the dye, and the labelled protein was separated from free dye and urea by passing the reaction mixture through a PD 10 column. The extent of labelling was checked by absorbance measurements using a value for ϵ_{337} of 4500 $M^{-1} cm^{-1}$ at 337 nm for 1,5-IAEDANS⁸⁵ and a value for ϵ_{360} of 13,300 $M^{-1} cm^{-1}$ at 360 nm for acrylodan.⁸⁶

Spectrophotometric measurements

Absorption spectra were recorded on a Shimadzu UV-2100 spectrophotometer using a cell of 1 cm path length. The protein concentration used for the measurements was $\sim 10 \mu$ M under various unfolding conditions.

Fluorescence measurements and data analysis

All the steady-state fluorescence measurements on TNB-labelled barstar mutants were carried out using a SPEX fluorolog (T-format) FL111 spectrofluorimeter by exciting Trp53 at 295 nm.

Time-resolved fluorescence intensity and anisotropy decay measurements were carried out using a time-correlated single photon counting setup. For TR-FRET measurements on TNB-labelled barstar mutants, 1 ps pulses of 887 nm radiation from the Ti-sapphire femto/pico second (Spectra Physics, Mountain View, CA) laser, pumped by an Nd:YLF laser (Millenia X, Spectra Physics), were frequency tripled to 295 nm by using a frequency doubler/tripler (GWU, Spectra physics). Similarly for anisotropy measurements on 1,5-IAEDANS or acrylodan labelled barstar mutants, 1 ps pulses of 830 nm radiation were frequency doubled to 415 nm. Fluorescence decay curves were obtained at the laser repetition rate of 4 MHz by a micro-channel plate photomultiplier (model R2809u; Hamamatsu Corp.) coupled to a time-correlated-single-photon-counting setup. The instrument response functions (IRF) at 295 nm and 415 nm were obtained using a dilute colloidal suspension of dried non-dairy coffee whitener. The width (FWHM) of the IRF was ~ 40 ps. Fluorescence emission measurements from TNB-labelled samples, excited at 295 nm, were done at 380 nm by using a combination of a monochromator and a 320 nm cut-off filter. 1,5-IAEDANS or acrylodan-labelled samples were excited at 415 nm and their emission measurements were done at 500 nm and 510 nm, respectively, by using proper excitation wavelength cut-off filters. In fluorescence lifetime measurements, the emission was monitored at the magic angle (54.7°) to eliminate the contribution from the decay of anisotropy. In time-resolved anisotropy measurements, the emission was collected at directions parallel (I_{\parallel}) and perpendicular (I_{\perp}) to the polarization of the excitation beam. The anisotropy was calculated as:

$$r(t) = \frac{I_{\parallel}(t) - I_{\perp}(t)G(\lambda)}{I_{\parallel}(t) + 2I_{\perp}(t)G(\lambda)} \quad (1)$$

where $G(\lambda)$ is the geometry factor at the wavelength λ of emission. The G factor of the emission collection optics was determined in separate experiments using a standard sample (NATA for TR-FRET measurements on TNB-labelled samples; and fluorescein for anisotropy decay measurements on 1,5-IAEDANS or acrylodan-labelled samples). The fluorescence decay curves at the magic angle were analysed by deconvoluting the observed decay with the IRF to obtain the intensity decay function represented as a sum of three or four exponentials:

$$I(t) = \sum \alpha_i \exp(-t/\tau_i) \quad i = 1-4 \quad (2)$$

where $I(t)$ is the fluorescence intensity at time t and α_i is the amplitude of the i th lifetime τ_i such that $\sum_i \alpha_i = 1$. Time-resolved anisotropy decay curves were analysed based on the model:

$$I_{\parallel}(t) = I(t)[1 + 2r(t)]/3 \quad (3)$$

$$I_{\perp}(t) = I(t)[1 - r(t)]/3 \quad (4)$$

$$r(t) = r_0\{\beta_1 \exp(-t/\phi_1) + \beta_2 \exp(-t/\phi_2)\} \quad (5)$$

where r_0 is the initial anisotropy. The values of r_0 for tryptophan, 1,5-IAEDANS and acrylodan were 0.20 ± 0.01 (at λ_{ex} 295 nm), 0.37 ± 0.01 (at λ_{ex} 415 nm) and 0.32 ± 0.01 (at λ_{ex} 415 nm), respectively. β_i is the amplitude associated with the i th rotational correlation time ϕ_i , such that $\sum_i \beta_i = 1$.

Distance calculation from TR-FRET measurements

The apparent mean distance (R) between the donor and acceptor pair in a system can be determined from the mean efficiency of energy transfer (E) between them, which in turn, is related to the fluorescence lifetimes of the donor, τ_{DA} and τ_{D} , in the presence and in the absence of the acceptor, respectively.

τ_{DA} and τ_{D} used in our estimations correspond to the mean lifetime ($= \sum_i \alpha_i \tau_i$ where i represents the lifetime components obtained from discrete analysis by equation (2) values:

$$E = \left(\frac{R_0^6}{R_0^6 + R^6} \right) = \left(1 - \frac{\tau_{\text{DA}}}{\tau_{\text{D}}} \right) \quad (6)$$

$$R_0 = 0.211[\kappa^2 n^{-4} Q_{\text{D}} J(\lambda)]^{1/6} \text{ \AA} \quad (7)$$

$$J(\lambda) = \frac{\int_0^{\infty} F_{\text{D}}(\lambda) \varepsilon_{\text{A}}(\lambda) \lambda^4 d\lambda}{\int_0^{\infty} F_{\text{D}}(\lambda) d\lambda} \text{ M}^{-1} \text{ cm}^{-1} (\text{nm}^4) \quad (8)$$

where R_0 is the distance between donor and acceptor at which the energy transfer efficiency is 50%, n is the refractive index of the medium. κ^2 is the orientation factor for the emission and absorption dipoles and its value (0–4) depends on their relative orientation. In our calculations, the value of κ^2 is taken to be 2/3, for random orientation of donor and acceptor with respect to each other, an assumption that is applicable to unfolded protein chains. Q_{D} is the quantum yield of the donor. Since Q_{D} is proportional to fluorescence lifetime, its value under various unfolding conditions was determined from the known Q_{D} value of 0.11 in 7 M urea.⁸⁷ $J(\lambda)$ is the

overlap integral, $F_{\text{D}}(\lambda)$ is the fluorescence intensity of the donor in the absence of acceptor, $\varepsilon_{\text{A}}(\lambda)$ is molar extinction coefficient of the acceptor. The overlap integral $J(\lambda)$ under the various denaturant conditions was determined from equation (8) by recording the emission spectrum of unlabelled protein ($\lambda_{\text{ex}} = 295$ nm) and the molar extinction coefficient spectrum of TNB-labelled proteins under the same conditions. Changes in the values of the parameters n^{-4} , Q_{D} and $J(\lambda)$ from one unfolding condition to another unfolding condition are very small. Also, since R_0 has a 1/6th power dependence on n^{-4} , Q_{D} and $J(\lambda)$, small changes in the values of these parameters under any unfolding conditions in the unfolded protein-baseline region do not result in a significant change of R_0 , whose value under any unfolding condition is $22.2(\pm 0.1)$ \AA. This value is the same as reported earlier.⁸⁷

Maximum entropy method of analysis (MEM)

The lifetime distributions in the unfolded form of barstar were obtained by analysing the fluorescence decay kinetics of Trp53 using MEM analysis.^{36,46–48} MEM analysis treats the fluorescence decay as arising from a distribution of a large number (~100–150) of discrete lifetime values equally spaced in the $\log(\tau)$ space covering the range from 10 ps to 10 ns, or in a similar range depending on the nature of the fluorescent molecule. The analysis begins by assigning equal probability (amplitude) to all the lifetime values. Subsequently, in each iteration during the analysis, the distribution is modified leading to minimization of χ^2 and maximization of the Shannon–Jaynes entropy function, $S = -\sum_i p_i \log p_i$, where $p_i = \alpha_i / \sum \alpha_i$ is the probability (amplitude) of i th lifetime. If many possible distributions have the same or similar value of χ^2 , then the maximum entropy criterion selects the distribution for which S is maximum. Thus, MEM analysis results in a lifetime distribution that is independent of any mathematical model. The values of χ^2 were in the range of 1.0 to 1.05 for all the MEM analyses. Peak values in the MEM distribution agreed with those obtained from discrete lifetime analyses, within 5%.

Acknowledgements

We thank Kalyan Sinha for donating some of the mutant protein preparations used in this study, as well as for critical comments on the manuscript; and Professor N. Periasamy for providing us with the home-developed software used in the analysis of the time-resolved fluorescence data. A.M.S. acknowledges receipt of a Kanwal Rekhi career development fellowship from the TIFR Endowment Fund. This work was funded by the Tata Institute of Fundamental Research, and by the Department of Science and Technology, Government of India.

References

1. Evans, P. A., Topping, K. D., Woolfson, D. N. & Dobson, C. M. (1991). Hydrophobic clustering in

- nonnative states of a protein: interpretation of chemical shifts in NMR spectra of unfolded states of lysozyme. *Proteins: Struct. Funct. Genet.* **9**, 248–266.
- Neri, D., Billeter, M., Wider, G. & Wuthrich, K. (1992). NMR determination of residual structure in a urea-unfolded protein, the 434-repressor. *Science*, **257**, 1559–1563.
 - Schwarzinger, S., Wright, P. E. & Dyson, H. J. (2002). Molecular hinges in protein folding: the urea-unfolded state of apomyoglobin. *Biochemistry*, **41**, 12681–12686.
 - Logan, T. M., Theriault, Y. & Fesik, S. W. (1994). Structural characterization of the FK506 binding protein unfolded in urea and guanidine hydrochloride. *J. Mol. Biol.* **236**, 637–648.
 - Garcia, P., Merola, F., Receveur, V., Blandin, P., Minard, P. & Desmadril, M. (1998). Steady state and time-resolved fluorescence study of residual structures in an unfolded form of yeast phosphoglycerate kinase. *Biochemistry*, **37**, 7444–7455.
 - Fersht, A. R. (1995). Optimization of rates of protein folding: the nucleation–condensation mechanism and its implications. *Proc. Natl Acad. Sci. USA*, **92**, 10869–10873.
 - Sosnick, T., 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.
 - Ferguson, N., Day, R., Johnson, C. M., Allen, M. D., Daggett, V. & Fersht, A. R. (2005). Simulation and experiment at high temperatures: ultrafast folding of a thermophilic protein by nucleation–condensation. *J. Mol. Biol.* **347**, 855–870.
 - Uversky, V. N., Gillespie, J. R. & Fink, A. L. (2000). Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins: Struct. Funct. Genet.* **41**, 415–427.
 - Uversky, V. N. (2002). What does it mean to be natively unfolded? *Eur. J. Biochem.* **269**, 2–12.
 - Wright, P. E. & Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure–function paradigm. *J. Mol. Biol.* **293**, 321–331.
 - Dobson, C. M. (2001). The structural basis of protein folding and its links with human disease. *Phil. Trans. Roy. Soc. ser. B*, **356**, 133–145.
 - Kelly, J. W. (2002). Towards an understanding of amyloidogenesis. *Nature Struct. Biol.* **9**, 323–325.
 - Zerovnik, E. (2002). Amyloid-fibril formation. Proposed mechanisms and relevance to conformational disease. *Eur. J. Biochem.* **269**, 3362–3371.
 - Dyson, H. J. & Wright, P. E. (2002). Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. *Advan. Protein Chem.* **62**, 311–340.
 - Wong, K. B., Clarke, J., Bond, C. J., Neira, J. L., Freund, S. M., Fersht, A. R. & Daggett, V. (2000). Towards a complete description of the structural and dynamic properties of the unfolded state of barnase and the role of residual structure in folding. *J. Mol. Biol.* **296**, 1257–1282.
 - Oliveberg, M., Vuilleumier, S. & Fersht, A. R. (1994). Thermodynamic study of the acid denaturation of barnase and its dependence on ionic strength: evidence for residual electrostatic interactions in the acid/thermally unfolded state. *Biochemistry*, **33**, 8826–8832.
 - Oliveberg, M., Arcus, V. L. & Fersht, A. R. (1995). pK_a values of carboxyl groups in the native and unfolded states of barnase: the pK_a values of the unfolded state are on average 0.4 unit lower than those of model compounds. *Biochemistry*, **34**, 9424–9433.
 - Ackerman, M. S. & Shortle, D. (2002). Robustness of the long-range structure in unfolded staphylococcal nuclease to changes in amino acid sequence. *Biochemistry*, **41**, 13791–13797.
 - Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T. *et al.* (2002). Long-range interactions within a nonnative protein. *Science*, **295**, 1719–1722.
 - Sosnick, T. R. & Trehwella, J. (1992). Unfolded states of ribonuclease A have compact dimensions and residual secondary structure. *Biochemistry*, **31**, 8329–8335.
 - Flanagan, J. M., Kataoka, M., Fujisawa, T. & Engelman, D. M. (1993). Mutations can cause large changes in the conformation of a unfolded protein. *Biochemistry*, **32**, 10359–10370.
 - Smith, L. J., Fiebig, K. M., Schwalbe, H. & Dobson, C. M. (1996). The concept of a random coil. Residual structure in peptides and unfolded proteins. *Fold. Des.* **1**, R95–R106.
 - Chatterjee, A., Mridula, P., Mishra, R. K., Mittal, R. & Hosur, R. V. (2005). Folding regulates autoprocessing of HIV-1 protease precursor. *J. Biol. Chem.* **280**, 11369–11378.
 - 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.
 - 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.
 - Buscaglia, M., Kubelka, J., Eaton, W. A. & Hofrichter, J. (2005). Determination of ultrafast protein folding rates from loop formation dynamics. *J. Mol. Biol.* **347**, 657–664.
 - Swaminathan, R., Nath, U., Udgaonkar, J. B., Periasamy, N. & Krishnamoorthy, G. (1996). Motional dynamics of a buried tryptophan reveals the presence of partially structured forms during denaturation of barstar. *Biochemistry*, **35**, 9150–9157.
 - Fa, M., Karolin, J., Aleshkov, S., Strandberg, L., Johansson, L. B. & Ny, T. (1995). Time-resolved polarized fluorescence spectroscopy studies of plasminogen activator inhibitor type 1: conformational changes of the reactive center upon interactions with target proteases, vitronectin and heparin. *Biochemistry*, **34**, 13833–13840.
 - Broos, J., Visser, A. J. W. G., Engbersen, J. F. J., Verboom, W., van Hoek, A. & Reinhoudt, D. N. (1995). Flexibility of enzymes suspended in organic solvents probed by time-resolved fluorescence anisotropy. Evidence that enzyme activity and enantioselectivity are directly related to enzyme flexibility. *J. Am. Chem. Soc.* **117**, 12657–12663.
 - Navon, A., Ittah, V., Landsman, P., Scheraga, H. A. & Haas, E. (2001). Distributions of intramolecular distances in the reduced and unfolded states of bovine pancreatic ribonuclease A. Folding initiation structures in the C-terminal portions of the reduced protein. *Biochemistry*, **40**, 105–118.
 - Sridevi, K., Lakshmikanth, G. S., Krishnamoorthy, G. & Udgaonkar, J. B. (2004). Increasing stability reduces conformational heterogeneity in a protein folding intermediate ensemble. *J. Mol. Biol.* **337**, 699–711.
 - 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.
34. Lakowicz, J. R., Gryczynski, I., Laczko, G., Wiczko, W. & Johnson, M. L. (1994). Distribution of distances between the tryptophan and the N-terminal residue of melittin in its complex with calmodulin, troponin C, and phospholipids. *Protein Sci.* **3**, 628–637.
 35. Wu, P. & Brand, L. (1994). Resonance energy transfer: methods and applications. *Anal. Biochem.* **218**, 1–13.
 36. 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.
 37. Schuler, B., Lipman, E. A., Steinbach, P. J., Kumke, M. & Eaton, W. A. (2005). Polyproline and the “spectroscopic ruler” revisited with single-molecule fluorescence. *Proc. Natl Acad. Sci. USA*, **102**, 2754–2759.
 38. 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.
 39. McCarney, E. R., Werner, J. H., Bernstein, S. L., Ruczinski, I., Makarov, D. E., Goodwin, P. M. & Plaxco, K. W. (2005). Site-specific dimensions across a highly denatured protein; a single molecule study. *J. Mol. Biol.* **352**, 672–682.
 40. Pletneva, E. V., Gray, H. B. & Winkler, J. R. (2005). Many faces of the unfolded state: conformational heterogeneity in unfolded yeast cytochrome *c*. *J. Mol. Biol.* **345**, 855–867.
 41. Sinha, K. K. & Udgaonkar, J. B. (2005). Dependence of the size of the initially collapsed form during the refolding of barstar on denaturant concentration: evidence for a continuous transition. *J. Mol. Biol.* **353**, 704–718.
 42. Bhavesh, N. S., Juneja, J., Udgaonkar, J. B. & Hosur, R. V. (2004). Native and nonnative conformational preferences in the urea-unfolded state of barstar. *Protein Sci.* **13**, 3085–3091.
 43. Pradeep, L. & Udgaonkar, J. B. (2004). Effect of salt on the urea-unfolded form of barstar probed by *m* value measurements. *Biochemistry*, **43**, 11393–11402.
 44. Lakowicz, J. R., Gryczynski, I., Cheung, H. C., Wang, C. K., Johnson, M. L. & Joshi, N. (1988). Distance distributions in proteins recovered by using frequency-domain fluorometry. Applications to troponin I and its complex with troponin C. *Biochemistry*, **27**, 9149–9160.
 45. Haas, E., Katchalski-Katzir, E. & Steinberg, I. Z. (1978). Brownian motion of the ends of oligopeptide chains in solution as estimated by energy transfer between the chain ends. *Biopolymers*, **17**, 11–31.
 46. Brochon, J. C. (1994). Maximum entropy method of data analysis in time-resolved spectroscopy. *Methods Enzymol.* **240**, 262–311.
 47. 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.
 48. 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.
 49. Lakshmikanth, G. S. & Krishnamoorthy, G. (1999). Solvent-exposed tryptophans probe the dynamics at protein surfaces. *Biophys. J.* **77**, 1100–1106.
 50. Tanford, C., Kawahara, K. & Lapanje, S. (1966). Proteins in 6-M guanidine hydrochloride. Demonstration of random coil behavior. *J. Biol. Chem.* **241**, 1921–1923.
 51. Tanford, C. (1968). Protein denaturation. Part A. Characterization of the unfolded state. *Advan. Protein Chem.* **23**, 121–217.
 52. Dill, K. A. & Shortle, D. (1991). Unfolded states of proteins. *Annu. Rev. Biochem.* **60**, 795–825.
 53. Baldwin, R. L. (2002). A new perspective on unfolded proteins. *Advan. Protein Chem.* **62**, 361–367.
 54. Shortle, D. (1996). The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J.* **10**, 27–34.
 55. Millett, I. S., Doniach, S. & Plaxco, K. W. (2002). Toward a taxonomy of the denatured state: small angle scattering studies of unfolded proteins. *Advan. Protein Chem.* **62**, 241–262.
 56. Rose, G. D., ed. (2002). *Unfolded Proteins Advances in Protein Chemistry*, vol. 62, Academic Press, San Diego, CA.
 57. Fitzkee, N. C. & Rose, G. D. (2004). Reassessing random-coil statistics in unfolded proteins. *Proc. Natl Acad. Sci. USA*, **101**, 12497–12502.
 58. Chan, H. S. & Dill, K. A. (1991). Polymer principles in protein structure and stability. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 447–490.
 59. Alonso, D. O. & Dill, K. A. (1991). Solvent denaturation and stabilization of globular proteins. *Biochemistry*, **30**, 5974–5985.
 60. Moglich, A., Krieger, F. & Kiefhaber, T. (2005). Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains. *J. Mol. Biol.* **345**, 153–162.
 61. 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.
 62. Kuzmenkina, E. V., Heyes, C. D. & Nienhaus, G. U. (2005). Single-molecule Förster resonance energy transfer study of protein dynamics under denaturing conditions. *Proc. Natl Acad. Sci. USA*, **102**, 15471–15476.
 63. Kohn, J. E., Millett, I. S., Jacob, J., Zagrovic, B., Dillon, T. M., Cingel, N. *et al.* (2004). Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl Acad. Sci. USA*, **101**, 12491–12496.
 64. Panick, G., Malessa, R., Winter, R., Rapp, G., Frye, K. J. & Royer, C. A. (1998). Structural characterization of the pressure-denatured state and unfolding/refolding kinetics of staphylococcal nuclease by synchrotron small-angle X-ray scattering and Fourier-transform infrared spectroscopy. *J. Mol. Biol.* **275**, 389–402.
 65. Millett, I. S., Townsley, L. E., Chiti, F., Doniach, S. & Plaxco, K. W. (2002). Equilibrium collapse and the kinetic “foldability” of proteins. *Biochemistry*, **41**, 321–325.
 66. Kataoka, M., Nishii, I., Fujisawa, T., Ueki, T., Tokunaga, F. & Goto, Y. (1995). Structural characterization of the molten globule and native states of apomyoglobin by solution X-ray scattering. *J. Mol. Biol.* **249**, 215–228.
 67. Zhou, J. M., Fan, Y. X., Kihara, H., Kimura, K. & Amemiya, Y. (1997). Unfolding of dimeric creatine

- kinase in urea and guanidine hydrochloride as measured using small angle X-ray scattering with synchrotron radiation. *FEBS Letters*, **415**, 183–185.
68. Chan, C. K., Hu, Y., 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.
69. Smith, C. K., Bu, Z., Anderson, K. S., Sturtevant, J. M., Engelman, D. M. & Regan, L. (1996). Surface point mutations that significantly alter the structure and stability of a protein's unfolded state. *Protein Sci.* **5**, 2009–2019.
70. Wandschneider, E. & Bowler, B. E. (2004). Conformational properties of the iso-1-cytochrome *c* unfolded state: dependence on guanidine hydrochloride concentration. *J. Mol. Biol.* **339**, 185–197.
71. 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.
72. Makhatazde, G. I. & Privalov, P. L. (1992). Protein interactions with urea and guanidinium chloride. A calorimetric study. *J. Mol. Biol.* **226**, 491–505.
73. 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.
74. Flory, P. J. (1953). *Principles of Polymer Chemistry*, Cornell University Press, Ithaca, NY.
75. Goldenberg, D. P. (2003). Computational simulation of the statistical properties of unfolded proteins. *J. Mol. Biol.* **326**, 1615–1633.
76. 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.
77. Lapidus, L. J., Steinbach, P. J., Eaton, W. A., Szabo, A. & Hofrichter, J. (2002). Effects of chain stiffness on the dynamics of Loop formation in polypeptides. Appendix: testing a 1-dimensional diffusion model for peptide dynamics. *J. Phys. Chem. B*, **106**, 11628–11640.
78. Hagen, S. J., Hofrichter, J. & Eaton, W. A. (1997). Rate of intrachain diffusion of unfolded cytochrome *c*. *J. Phys. Chem. B*, **101**, 2352–2365.
79. Wong, K. B., Freund, S. M. & Fersht, A. R. (1996). Cold denaturation of barstar: ^1H , ^{15}N and ^{13}C NMR assignment and characterisation of residual structure. *J. Mol. Biol.* **259**, 805–818.
80. Tezcan, F. A., Findley, W. M., Crane, B. R., Ross, S. A., Lyubovitsky, J. G., Gray, H. B. & Winkler, J. R. (2002). Using deeply trapped intermediates to map the cytochrome *c* folding landscape. *Proc. Natl Acad. Sci. USA*, **99**, 8626–8630.
81. Lyubovitsky, J. G., Gray, H. B. & Winkler, J. R. (2002). Mapping the cytochrome *c* folding landscape. *J. Am. Chem. Soc.* **124**, 5481–5485.
82. Khurana, R., Hate, A. T., Nath, U. & Udgaonkar, J. B. (1995). pH dependence of the stability of barstar to chemical and thermal denaturation. *Protein Sci.* **4**, 1133–1144.
83. Nath, U. & Udgaonkar, J. B. (1997). Folding of tryptophan mutants of barstar: evidence for an initial hydrophobic collapse on the folding pathway. *Biochemistry*, **36**, 8602–8610.
84. Ramachandran, S. & Udgaonkar, J. B. (1996). Stabilization of barstar by chemical modification of the buried cysteines. *Biochemistry*, **35**, 8776–8785.
85. Lillo, M. P., Beechem, J. M., Szpikowska, B. K., Sherman, M. A. & Mas, M. T. (1997). 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.
86. Post, P. L., Trybus, K. M. & Taylor, D. L. (1994). A genetically engineered, protein-based optical biosensor of myosin II regulatory light chain phosphorylation. *J. Biol. Chem.* **269**, 12880–12887.
87. 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.

Edited by C. R. Matthews

(Received 12 November 2005; received in revised form 2 March 2006; accepted 8 March 2006)
Available online 24 March 2006