PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2018, **20**, 3216

Site-specific time-resolved FRET reveals local variations in the unfolding mechanism in an apparently two-state protein unfolding transition†

Sandhya Bhatia,[a] G. Krishnamoorthy[b] and Jayant B. Udgaonkar*[a]

Received 12th September 2017,
Accepted 16th October 2017

DOI: 10.1039/c7cp06214a

rsc.li/pccp

Protein folding/unfolding transitions between the native (N) and unfolded (U) states are usually describable as two-state, only because of the dominant presence of the N and/or U states, because of which high energy intermediates remain undetected. Delineation of the cooperativity underlying these transitions, and characterization of high energy intermediates that are populated sparsely, have been difficult challenges, especially under equilibrium conditions, and require the use of a sensitive probe that reports on both the structures and population distributions of the partially unfolded intermediates. In this study, the use of multisite time-resolved FRET to monitor structural change in five specific segments of the small protein monellin, has brought out local deviations from two-state behavior during unfolding. It is shown that in some segments of the protein structure, denaturant-induced unfolding proceeds first by gradual expansion of the N state, then by an all-or-none transition from the N state ensemble to the U state ensemble, followed finally by expansion of the U state. Segments encompassing the sole helix appear, however, to unfold completely through a gradual transition from the N to U states. Finally, it is shown that equilibrium unfolding of monellin is not only heterogeneous, but that the degree of non-cooperativity differs between the sole α-helix and different parts of the β-sheet.

Introduction

The extent of cooperativity in protein folding and unfolding reactions has been difficult to characterize because the reactions are usually monitored by ensemble-averaging probes.[1–5] When monitored by such probes, the unfolding of many proteins under equilibrium conditions appears to occur in a 'two-state' manner, because the probes are typically not sensitive enough to distinguish low populations of partially unfolded intermediates in the dominating presence of the native (N) and unfolded (U) states. The limited cooperativity of folding/unfolding reactions becomes evident only in kinetic experiments,[6–15] although not always, as the thermodynamic and kinetic criteria of two-state folding[16,17] are met even when high energy intermediates are present.[18,19] On the other hand, for several proteins, protein folding reactions have been shown to be completely non-cooperative (gradual).[3,4,20–26]

Protein unfolding reactions are not really expected to be two-state as they might seem to be when probed by ensemble-averaging probes. It is, however, difficult to determine, where in the structure deviations from two-state behavior occur. Deviations could be of two main types: (1) different structural segments may have different stabilities that are dispersed around the global stability.[27–32] (2) Some of the structural segments may not unfold in an all-or-none manner but in a gradual manner,[21,22,25,33–39] but the non-cooperativity may be hidden underneath the apparent cooperativity observed when the transition is measured by an ensemble-averaging probe. Thus, it becomes important not only to determine the extent of cooperativity present in protein folding reactions, but also to localize the observed cooperativity to specific regions in the protein structure, in order to understand the structural and energetic origins of folding cooperativity.

Of particular importance is whether all secondary structures can form synchronously in one step or not. ϕ-Value analysis of the folding of many proteins has suggested that secondary structures are only partially formed in the transition state perched atop the dominant free energy barrier that separates the U and N states.[40,41] It has been difficult to determine whether individual secondary structures form cooperatively or non-cooperatively, in the context of the overall protein structure. It is not known how the cooperativity of helix formation is modulated by the structural and energetic constraints placed on the helical segment by the rest of the protein sequence and structure.[4,25,42]

[a] *National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru 560065, India. E-mail: jayant@ncbs.res.in; Fax: +91-80-23636662; Tel: +91-80-23666150*

[b] *Department of Biotechnology, Anna University, Chennai 600025, India*

† Electronic supplementary information (ESI) available: ESI contains (1) Fig. S1–S11 and (2) Tables S1–S4. See DOI: 10.1039/c7cp06214a

Similarly, the cooperativity of assembly of β-strands to form a β-sheet in a protein remains poorly understood.[43–47] It is important to determine how structural and sequence constraints dictated by the entire protein structure, affect the assembly of various secondary structural units within a protein.

The best way to investigate folding cooperativity is to determine the population distribution of a particular structural property. For example, both single molecule FRET (sm-FRET)[48,49] and time-resolved FRET (tr-FRET)[35,50] measurements can reveal how an intra-molecular distance distribution changes during protein folding/unfolding, and whether the change occurs cooperatively or non-cooperatively. In such measurements, it is important to monitor multiple intra-molecular distances, because some distances may change in an all-or-none manner, and other distances may change continuously. Hence, different segments of the sequence/structure may fold/unfold either cooperatively or gradually.[4,22,24,51,52] The sm-FRET and tr-FRET methodologies can therefore be useful for investigating the local cooperativity of folding/unfolding of different secondary structural units in a protein.

Monellin is a small 97-residue protein, in which a single α-helix is packed against a 5-stranded β-sheet. Its folding mechanisms have been studied in both its naturally occurring two-chain form[53,54] and its artificially created single-chain form (MNEI).[4,12,15,19,22,25,55–60] Both forms have been shown to fold *via* multiple pathways.[12,54,57] The single chain form, MNEI, folds *via* an early kinetic molten globule,[60] that itself forms in multiple steps.[15] Kinetic FRET measurements have revealed that the unfolding of MNEI begins with the formation of a dry molten globule,[56] after which the structure is lost as the protein molecule swells in a gradual manner.[22] Native-state hydrogen exchange (HX) measurements indicate that unfolding is gradual in the absence of denaturant, but becomes cooperative in the presence of GdnHCl,[25] and that the cooperativity originates from the β2–β3 strands.[4] But HX methods, in general, are restricted to regions of the protein that are structured and buried, and usually provide no information on more than 60% of the sequence that offers little or no protection against HX. Hence, little is known about the stabilities of different structural elements, especially the secondary structural units, and whether these structural elements unfold in a cooperative or a non-cooperative manner.

It is also important to note that the complexity of the folding and unfolding of MNEI has been revealed only in kinetic experiments. Under equilibrium conditions, transient intermediates are too unstable to be populated sufficiently to be detected. Hence, the equilibrium unfolding data are describable by a two state N ↔ U model for unfolding. MNEI is therefore a good candidate protein for demonstrating that when a suitable site-specific high resolution structural method is used, even equilibrium studies can reveal structurally localized deviations from two state unfolding, which are otherwise concealed when an ensemble-averaging probe is utilized.

In this study, the thermodynamic cooperativity of unfolding of various structural regions of MNEI has been probed by site-specific time-resolved FRET measurements. Equilibrium unfolding has been monitored using five FRET pairs, encompassing different parts of the α-helix, the β-sheet, the helix-core interface, and the end to end distance of MNEI. The results obtained by coupling model-free MEM analysis to tr-FRET data suggest that MNEI unfolds overall in a non-cooperative manner, but the degree of cooperativity varies depending upon the structural region monitored. Distinct regions of the protein could be classified into two broad categories, based on the nature of unfolding: (a) segments where the process is fully non-cooperative even in the presence of denaturant; this was observed for the unfolding of the helix; (b) segments where the unfolding involved both cooperative changes from the N to U state ensemble, as well as gradual changes accompanying the expansion of the N and U state ensemble; this was seen for the rest of the structural regions which all involve the β-sheet. Surprisingly, the degree of non-cooperativity within the β-sheet varied from one region to another.

Materials and methods

Protein expression, purification and TNB labeling

The wild-type protein (W4C42) has a single tryptophan (W4) residue in the first β-strand and a single buried cysteine (C42) in the second β-strand. The single-tryptophan-single-cysteine-containing mutant variants, C42AQ29C (W4C29), C42AP97C (W4C97), W4YF19W (W19C42) and C42AW4YF19WQ29C (W19C29), were generated by site-directed mutagenesis. The proteins were purified as described previously,[12] except that the second chromatographic separation was not carried out, as it was found unnecessary.

The single cysteine residue in the protein was conjugated covalently to the acceptor thionitrobenzoate (TNB) moiety. The labeling protocol was similar to that described earlier.[22] In brief, the unlabeled protein was unfolded in 4 M GdnHCl and incubated with DTNB solution at pH 8 for $\geq$2 h; this was followed by desalting to get rid of excess label. The purity of each protein was confirmed by electrospray ionization mass spectrometry (ESI-MS). The mass of the labeled protein showed the expected increase of 197 Da, corresponding to the mass of the TNB group, and the extent of labeling was >95%. The concentrations of the unlabeled protein solutions were determined from absorbance measurements, using the molar extinction coefficient value of 14 600 M^{-1} cm^{-1} for all the unlabeled proteins, except for W19C42 and W19C29, for which a value of 18 300 M^{-1} cm^{-1} was used (determined from measuring protein concentration using the BCA assay). For the TNB-labeled proteins, the contribution of the TNB absorbance to the total absorbance at 280 nm was taken to be 20%, as described earlier.[22]

Reagents

All the experiments were carried out at pH 8.0 and 25 °C. The reagents used in the experiments were of the highest purity grade from Sigma. Guanidine hydrochloride (GdnHCl) was purchased from USB, and was of the highest purity grade. Native buffer

contained 20 mM Tris and 250 μM EDTA, and unfolding buffer contained, in addition, GdnHCl. 1 mM DTT was used for all experiments with unlabeled proteins. The concentrations of GdnHCl solutions were determined by the measurements of the refractive index on an Abbe 3L refractometer from Milton Roy. All buffers and solutions were passed through 0.22 μm filters before use. A protein concentration of 10 μM was used for all experiments.

Measurement of fluorescence and far UV CD spectra

Fluorescence spectra were collected on a Spex Fluoromax 3 spectrofluorometer. The protein samples were excited at 295 nm, and emission was collected from 310 to 450 nm, with a response time of 1 s, and excitation and emission bandwidths of 0.5 and 5 nm, respectively. Each spectrum was recorded as the average of two fluorescence emission wavelength scans.

Far-UV CD spectra were acquired using a Jasco J815 spectropolarimeter using a 0.2 cm path length cuvette. The far-UV CD spectra were collected with a bandwidth of 1 nm, a response time of 1 s, and a scan speed of 20 nm min^{-1}. Fifteen wavelength scans were averaged for each spectrum.

Spectroscopic measurement of conformational stability

For equilibrium unfolding studies, the protein was incubated in different concentrations of GdnHCl (0 to 4 M) for $\geq$6 h prior to measurement. CD-monitored equilibrium unfolding experiments were carried out using a JASCO J-815 spectropolarimeter. The ellipticity at 222 nm was measured using a cuvette with a path length of 0.2 cm. The data were acquired using a bandwidth of 2 nm, and averaged for 60 s for each sample.

Steady state FRET measurements were carried out using the MOS-450 optical system (Biologic) using an excitation wavelength of 295 nm; emission was collected at 360 nm using a 10 nm band pass filter (Asahi spectra). The excitation slit width was kept at 2 nm. The data were acquired and averaged for 30 s for each sample.

Time resolved FRET monitored equilibrium unfolding experiments

Fluorescence lifetime decay curves were acquired for both the unlabeled and labeled proteins equilibrated with different GdnHCl concentrations (0 to 4 M) for $\geq$6 h prior to measurements.

For this purpose, 885 nm radiation from a femto second Ti-sapphire laser (Tsunami or Mai Tai HP, Spectra Physics) pumped by an Nd:YVO4 laser (Millennia X, Spectra Physics), pulsing at a basic repetition rate of $\sim$80 MHz, was used. The pulse width was $\sim$1 ps. The pulse repetition rate was decreased to 8 MHz using a pulse selector (model 3980, Spectra Physics). The fundamental 885 nm light was then frequency-tripled to 295 nm using a flexible second- and third-harmonic generator (GWU, Spectra Physics).

The resultant 295 nm radiation was used as the excitation source, and focused on to the sample in the cuvette after passing through a polarizer (Thorlabs) (selective to vertical polarization). The fluorescence emission was passed through a monochromator (H10-UV, Horiba Jobin-Yvon) set at 360 nm (bandwidth of 30 nm) onto a cooled micro channel plate PMT (model 3809U-50, Hamamatsu) biased at −3000 V, or on to an HPM-100-40 hybrid PMT detector operating at 83% gain voltage. A 325 nm long-pass filter was used to reduce contributions from scattered excitation light. A polarizer (Glan-Thompson), set at 54.7° (magic angle) with respect to excitation, was kept in front of the long-pass filter to avoid polarization effects.

Single-photon counting measurements were triggered by a photodiode (APD110, Thor Laboratories, or APM-400, Becker and Hickl) and measured using a TCSPC card (SPC-630, Becker-Hickl) *via* a preamplifier (Ortec). The time per channel was set to 48.8 ps for all measurements. The instrument response function (IRF) was recorded for each sample for deconvolution. This was measured using light scattered from a colloidal solution of non-dairy coffee whitener or LUDOX solution. The MCP-PMT consistently gave an IRF FWHM of $\sim$60 ps, and the HPM-100-40 detector gave an IRF FWHM of $\sim$120 ps. The intensity of the excitation pulse was attenuated using a neutral density filter.

The fluorescence lifetime of NATA (2.5–2.8 ns) in water was routinely recorded as a standard during each experiment. All decay transients were collected to a peak count of 20 000, and up to 99.9% of completion. The fluorescence intensity decay curves were recorded for both the unlabeled (donor-only protein) and the corresponding TNB-labeled protein (donor–acceptor protein), at increasing GdnHCl concentrations, on the same day. Measurements were taken in quick succession to reduce variations arising from instrumental instabilities.

Data analysis

Analysis of tr-FRET decay curves

Discrete analysis of fluorescence intensity decays. The fluorescence decay traces were deconvolved with the IRF, and analyzed as a sum of two, three or four exponentials:

$$F(t) = \sum_{i=1}^{n} \alpha_i e^{-\left(\frac{t}{\tau_i}\right)} \quad (1)$$

Here, α_i is the relative amplitude of the τ_i lifetime component, t is time and n ranges from 2 to 4, corresponding to the different numbers of exponentials required to fit the data.

The mean lifetime, τ_m was calculated as:

$$\tau_m = \frac{\sum \alpha_i \tau_i}{\sum \alpha_i} \quad (2)$$

$$\sum a_i = 1 \quad (3)$$

A fit was accepted if its χ^2 value was $\leq$1.3, and if the residuals were randomly distributed about 0.

MEM analysis of fluorescence intensity decays. Fluorescence intensity decays for the unlabeled and TNB-labeled mutant proteins were analyzed using a customized MEM analysis program.[61–64] MEM analysis is independent of any physical model or mathematical equation describing the distribution of lifetimes. Thus, it is different from other methods of analyzing fluorescence decays. The analysis begins with the assumption

that the decay originates from a distribution of fluorescence lifetimes in the range from 10 ps to 10 ns, or in a similar range, and that all lifetimes are equally probable (have equal amplitude). The distribution is modified in subsequent iterations of the analysis to result in a minimization of the residuals and χ^2 values, and maximization of the Shannon–Jaynes' entropy, S, which is described as

$$S = -\sum P_i \log(P_i) \quad (4)$$

P_i is the probability of the ith lifetime

$$P_i = \frac{\alpha_i}{\sum \alpha_i} \quad (5)$$

α_i is the amplitude of the ith lifetime.

Analysis parameters were optimized for obtaining precise and reproducible MEM distributions. The peak lifetime for each distribution refers to the lifetime corresponding to maximum amplitude in the lifetime distribution data.

The robustness of the lifetime distributions obtained by MEM analysis was checked exhaustively by collecting data on multiple samples of native protein, as well as of unfolded protein under the same conditions (data not shown). For each protein, each equilibrium unfolding experiment was repeated at least twice, so that MEM analyses were carried out on at least two data sets acquired from at least two independently prepared samples at each denaturation concentration. For each sample at any denaturant concentration, fluorescence decays were acquired at least twice, and each decay was analyzed by MEM to obtain a fluorescence lifetime distribution.

Values of χ^2 were in the range of 0.8 to 2 for all of the discrete and MEM analyses. The residuals were randomly distributed about zero.

Fitting of MEM analysis-derived distributions to a two-state model. A nonlinear, least-squares curve fitting algorithm written in MATLAB was used to fit the fluorescence lifetime distributions to a two-state model. The algorithm uses the spectra of the native, N(τ) and unfolded, U(τ) state fluorescence lifetime distributions as basis spectra, and performs a non-linear least-squares fit of all the fluorescence lifetime distributions at all GdnHCl concentrations to the following equation:

$$Y = f_N \times N(\tau) + f_U \times U(\tau) \quad (6)$$

f_N and f_U represent the fractions of native and unfolded protein, respectively, at each GdnHCl concentration, for the best possible fit to the two-state model. The mean amplitude of the two distributions obtained for each sample was used for this analysis. The root mean squared deviation (rmsd, Fig. S10, ESI†) was obtained at a particular GdnHCl concentration by taking a square root of the mean of the squares of residuals across all lifetime values.

FRET efficiency calculations

Time resolved FRET efficiency was calculated using the following equation:

$$E_{FRET} = 1 - \left(\frac{\tau_{DA}}{\tau_D}\right) \quad (7)$$

τ_{DA} and τ_D are the mean lifetimes (Fig. 4) of the unlabeled and TNB-labeled proteins, respectively.

The steady state FRET efficiency was calculated using the following formula:

$$E_{FRET} = 1 - \left(\frac{F_{DA}}{F_D}\right) \quad (8)$$

F_{DA} and F_D correspond to the concentration-normalized fluorescence intensity signals, of the unlabeled and TNB-labeled proteins, respectively.

Extraction of energy transfer parameters and determination of D–A distances

The distance, R, separating the donor and the acceptor, is related to the energy transfer efficiency, E, by the Förster equation:

$$R = R_0\left(\frac{1-E}{E}\right)^{\frac{1}{6}} \quad (9)$$

The value of R_0, the Förster distance, can be determined by using the equation:

$$R_0 = 0.211(Q_D J \kappa^2 n^{-4})^{1/6} \quad (10)$$

In eqn (10), Q_D is the quantum yield of donor fluorescence, J is the overlap integral, κ^2 is the orientation factor, and n is the refractive index of the medium.[65] These parameters can be determined experimentally. For the determination of the Q_D values of the N and U forms (Table S1, ESI†), the areas under the fluorescence emission spectra of different variants of MNEI (Fig. 2) were determined relative to the area of the spectrum of NATA whose quantum yield is known to be 0.14.[66] The overlap integral, J, is defined as the spectral overlap between the fluorescence emission spectrum of the donor and the absorbance spectrum of the acceptor, and is given by,

$$J = \frac{\int F(\lambda) \in (\lambda)\lambda^4 d\lambda}{\int F(\lambda) d\lambda} \left(M^{-1}\ cm^{-1}\ (nm)^4\right) \quad (11)$$

For the determination of J, the fluorescence emission spectra, $F(\lambda)$, of all the unlabeled proteins in their N and U states were collected as described above. Absorption spectra of all the TNB-labeled proteins in the N and U states were collected on a CARY100 double-beam spectrophotometer, with a bandwidth of 1 nm and a scan speed of 1 nm s^{-1}, using a cuvette with a path length of 1 cm. All absorption spectra were divided by the respective molar protein concentration to obtain $\varepsilon(\lambda)$. J was determined as the overlap between $F(\lambda)$ and $\varepsilon(\lambda)$ according to eqn (11), where the wavelength λ is in nm.[65] These values are listed in Table S3 (ESI†). The value of the refractive index, n, of the medium was determined to be 1.33 in 0 M GdnHCl, and 1.4 in 4 M GdnHCl. The value of the orientation factor, κ^2, depends on the relative orientation of the transition dipoles of the donor, W4 or W19, and the TNB acceptor. If the donor and acceptor are oriented randomly with respect to each other, the value of κ^2 is 2/3.[65,67] This condition is typically met when at least one of the two chromophores is rotating freely at rates faster than the fluorescence decay rate. The rotational dynamics of W4 and

fluorophore placed at each of the three acceptor (Cys) sites had been measured previously, by determination of the fluorescence anisotropy decay kinetics,[22,56] and had been shown to be faster than fluorescence decay. The rotational dynamics of W19 were found in the current study to be as fast as that of W4 (data not shown). Thus, using a value of 2/3 for κ^2 for all the FRET pairs appears reasonable. The observation that the intrinsic quantum yields of W4 and W19 in MNEI, in both the N and U states (Table S1, ESI†), are similar in value to that of free tryptophan,[66] also suggests that W4 and W19 rotate freely and do not interact with the rest of the protein.

The values of the Förster distance, R_0, and the D–A distances, R, for all the TNB-labeled proteins, are listed in Table S3 (ESI†). The calculated values of R_0 in the N and U states of any TNB-labeled protein are very similar and thus, it is justified to use an average R_0 value for a given FRET pair for determining distances across all denaturant concentrations.

Results

The structure and stability of the protein were not perturbed by mutation and labelling

In this study, five intra-molecular distances were monitored using five different single tryptophan, single cysteine containing mutant variants. The FRET donor was tryptophan, located at residue position 4 or 19, and the FRET acceptor was a TNB moiety attached to the single cysteine located at residue position 29, 42 or 97 (Fig. 1). The secondary structure was very similar to the different mutant variants and their labeled counterparts (Fig. S1, ESI†).

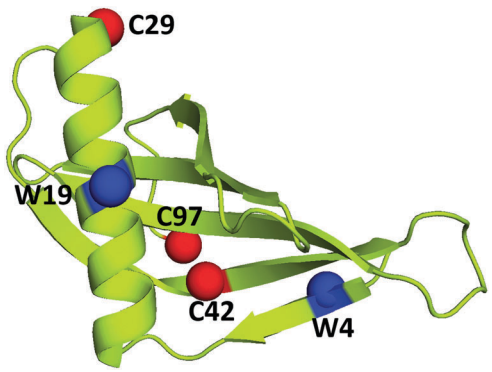

Fig. 1 Structure of single chain monellin (MNEI) showing the different residues used for monitoring FRET. Wild type MNEI has one tryptophan residue (W4) and one cysteine residue (C42). Cα atoms of tryptophan residues (used as the fluorescence donor) are shown as blue spheres, and the Cα atoms of cysteine residues (to which the fluorescence acceptor TNB was attached) are shown as red spheres. FRET was measured in single tryptophan-containing, single cysteine-containing variants. FRET changes monitored changes in the β-sheet (W4C42), at the helix-β-sheet interface (W19C42), in the end to end distance (W4C97), across the helix (W4C29), or across half the helix (W19C29). The distances between the center of the donor ring and Cβ atoms of the cysteine (attached to acceptor) in the N state, for the different FRET pairs monitored, are 13.0 Å (W4C42), 19.6 Å (W4C97), 16.1 Å (W19C42), 27.7 Å (W4C29) and 18.2 Å (W19C29). The structure has been drawn using PyMol (PDB ID: 1IV7).

The slight variation in the intensities of the far UV CD spectra of the mutant variants can be attributed to the contributions of Trp and Tyr to the far UV CD spectra being different when residue positions 4 and 19 differ in these aromatic residues. The stabilities of all the unlabeled proteins were not significantly different (<1 kcal mol^{-1}) from that of wt MNEI (W4C42) (Fig. S2 and Table S2, ESI†). This suggests that the mutations have a minimal perturbation on the protein structure, and hence it is unlikely that the global unfolding properties of the different variants are different from those of wt MNEI. Upon labeling, the stability was not affected significantly for W4C97, W4C29 and W19C29, whereas W4C42 and W19C42 were slightly destabilized. C42 is completely buried in the N state of MNEI (W4C42), and the modification of this residue, by the addition of the TNB moiety, may have disrupted packing interactions[68] and caused the decrease in stability. It should be noted that the slight destabilization caused by labeling or mutagenesis did not affect the global (un)folding properties of the protein: the unfolding/refolding kinetics of the W4C42-TNB, W4C97, W4C97-TNB, W4C29 and W4C29-TNB, were shown earlier to be not significantly different from that of wt MNEI (W4C42).[22,56,60] The refolding kinetics of the unlabeled and TNB-labeled variants of W19C42 and W19C29 were found in this study to also be similar to that of wt MNEI (data not shown).

Site-specific structural changes probed by tryptophan fluorescence and FRET

The folding–unfolding transition of MNEI was monitored by FRET at various concentrations of GdnHCl. The fluorescence emission maximum of W4, upon excitation at 295 nm, was at 346 nm for the native unlabeled protein, as seen for W4C42, W4C97 and W4C29 (Fig. 2). This suggested that W4 is partially buried (or surrounded by a non-polar environment) in the native state. Upon unfolding, the fluorescence emission maximum was at 356 nm (360 nm for W4C29), reflecting the exposure of W4 to the polar solvent. The fluorescence emission maximum of W19 at 350 nm in native W19C42 and W19C29 was found to be red-shifted with respect to the fluorescence emission spectrum of W4 in W4C42. This suggested that W19 is in a more polar environment than is W4 in W4C42. The red shift was observed for both the native and unfolded proteins. The fluorescence emission maximum of W4C42, W4C97, W4C29, W19C42 or W19C29 did not get affected by TNB labeling.

The quenching in fluorescence intensity upon labeling was observed at all wavelengths, and the extent of quenching was different for the N and U states. The extent of quenching indicated by the decrease in fluorescence intensity upon labeling was different for the different proteins, suggesting that quenching is distance dependent, and hence, due to FRET. The quenching was found to be independent of the protein concentration used, indicating the absence of intermolecular energy transfer (data not shown).

Dependence of the FRET efficiency on denaturant concentration appeared different across different structural segments

The equilibrium unfolding transitions monitored by steady-state FRET appeared to be two-state in nature, as indicated by

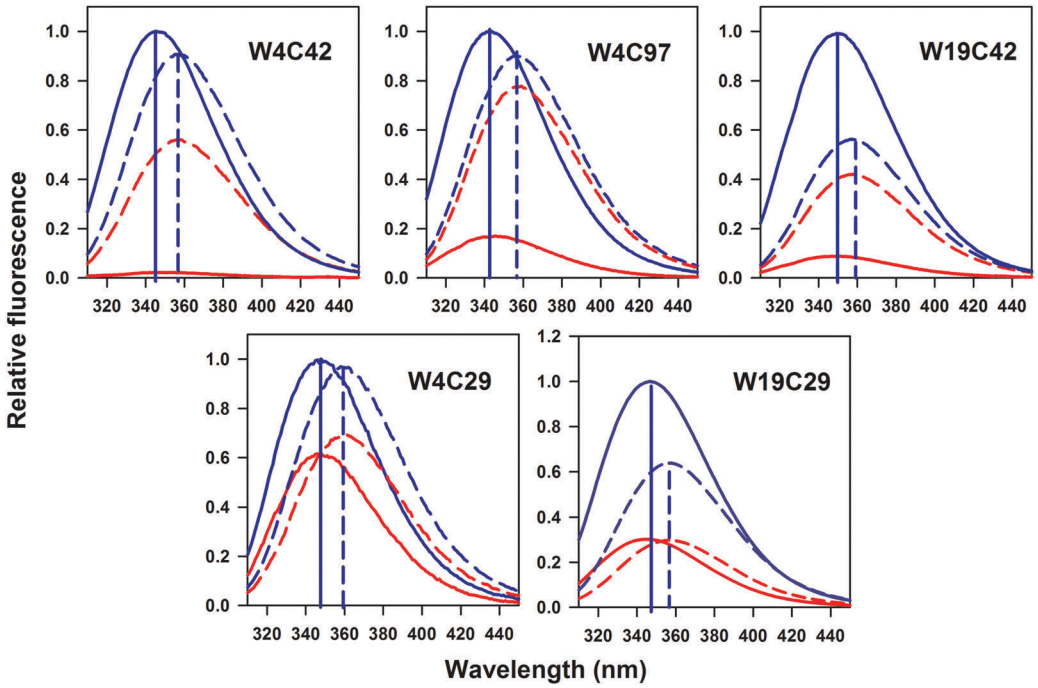

Fig. 2 Fluorescence emission spectra of the different mutant variants of MNEI. The excitation wavelength was 295 nm. The solid and dashed curves represent the spectra of the native and unfolded states, respectively, for the TNB-labeled (red) and unlabeled (blue) proteins. The solid and dashed vertical lines indicate the peak positions of the spectra of the native and unfolded unlabeled proteins, respectively.

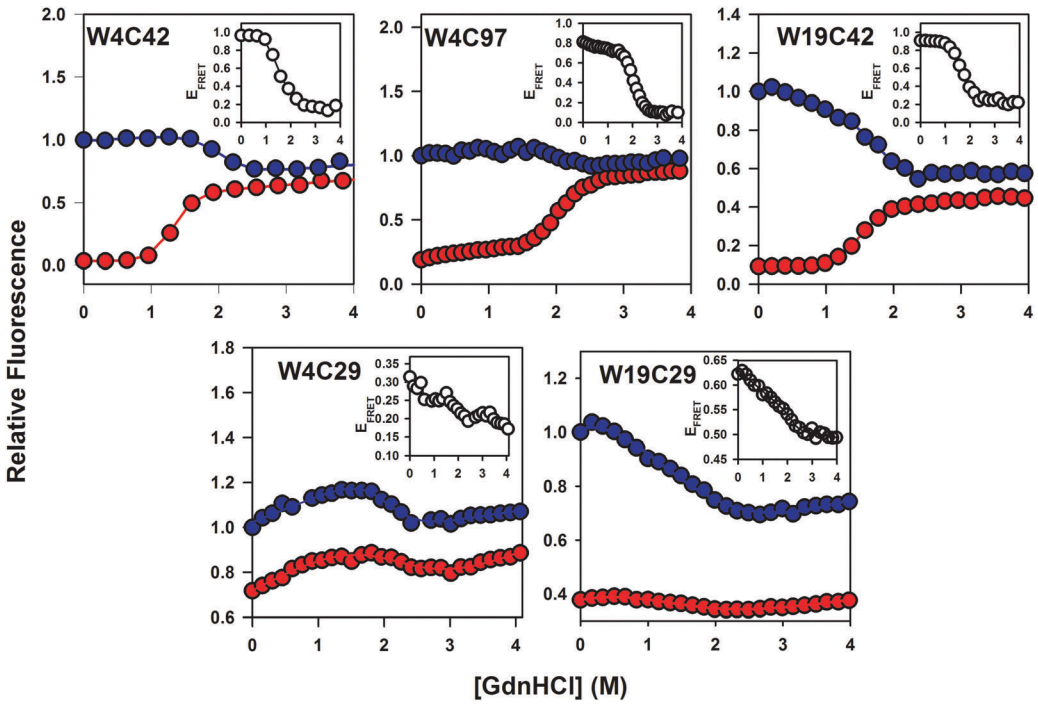

Fig. 3 Equilibrium unfolding transitions of MNEI monitored by steady state FRET. The fluorescence intensity was monitored at 360 nm, using an excitation wavelength of 295 nm. The blue and red circles correspond to the fluorescence intensities of the unlabeled and TNB-labeled proteins, respectively. The solid line passing through each dataset is a non-linear, least-squares fit to a two-state unfolding model.[69] The inset in each panel shows the dependence of the FRET efficiency on GdnHCl concentration. The solid line shown in the inset of panel W4C29 and W19C29 is drawn to guide the eye.

the sigmoidal dependence of the fluorescence intensity on GdnHCl concentration (Fig. 3). The sigmoidal equilibrium unfolding transitions for all the proteins (except for W19C42 and W19C29) were coincident when monitored by CD and by fluorescence of the tryptophan, again suggesting that all the proteins undergo two-state unfolding (Fig. S3, ESI†). However,

based on the dependence of the FRET efficiency on GdnHCl concentration (insets, Fig. 3), the equilibrium unfolding transitions of the proteins could be broadly divided into two categories. (a) In the case of W4C42, W4C97 and W19C42, the transitions monitored by FRET efficiency appeared highly cooperative, having a sigmoidal dependence on GdnHCl concentration, and (b) in the case of W4C29 and W19C29, the FRET efficiency did not have a sigmoidal dependence on GdnHCl concentration suggesting loss of cooperativity. In both cases, very broad transitions were observed; hence, the curves could not be modeled (Fig. 3, insets W4C29 and W19C29).

The dependence of mean fluorescence lifetime on GdnHCl concentration, obtained from discrete analysis of the time-resolved fluorescence data (Fig. 4), compared well with the steady-state fluorescence intensity data (Fig. 3). For W4C29, the mean lifetime data did not show any change in signal upon unfolding. This was expected as there was no significant change in the quantum yield of W4C29 for the unlabeled as well as the labeled proteins upon unfolding (Fig. 2 and Table S1, ESI†). It is important to note that the mean fluorescence lifetime is proportional to the quantum yield and is wavelength independent, unlike the fluorescence intensity. Hence, the dependence of fluorescence intensity (at 360 nm) on GdnHCl concentration seen for W4C29 and W4C29-TNB was not reproduced by mean fluorescence lifetime measurements.

For unlabeled W4C97, the mean lifetime data did not match the fluorescence intensity data (Fig. 3 and 4), possibly because of the presence of a very short lifetime component that was undetectable by the present TCSPC system, and which contributed significantly to the steady-state fluorescence intensity. Nonetheless, there was a good correspondence between the FRET efficiency of W4C97 obtained from fluorescence intensity and lifetime measurements (Fig. 3, inset W4C97 and Fig. 5). The two broad categories of FRET efficiency dependences on GdnHCl concentration (as described previously for fluorescence intensity measurements) were observed in the case of FRET efficiency and distances obtained from the fluorescence lifetime data as well (Fig. 5). However, at this stage, the two categories could not be classified as representing cooperative and continuous unfolding transitions, because of the inability of these measurements to resolve populations and differentiate between the two possibilities.

MEM analysis revealed the complex equilibrium unfolding transition of MNEI

Fluorescence intensity decay curves were analyzed by the Maximum Entropy Method (MEM) to generate fluorescence lifetime distributions (Fig. 6). In the case of W4C42-TNB, W4C97-TNB and W19C42-TNB, the fluorescence lifetime distributions obtained at various GdnHCl concentrations, were bimodal in nature. One distribution, encompassing lifetimes between 0.01 and 1 ns, appeared to originate from native-like (N-like) forms, and the other distribution, encompassing lifetimes >1 ns, appeared to originate from unfolded-like (U-like) forms.

The salient features of the data were: (a) in the N-like forms, a gradual shift of the peak of the distribution towards longer lifetimes,

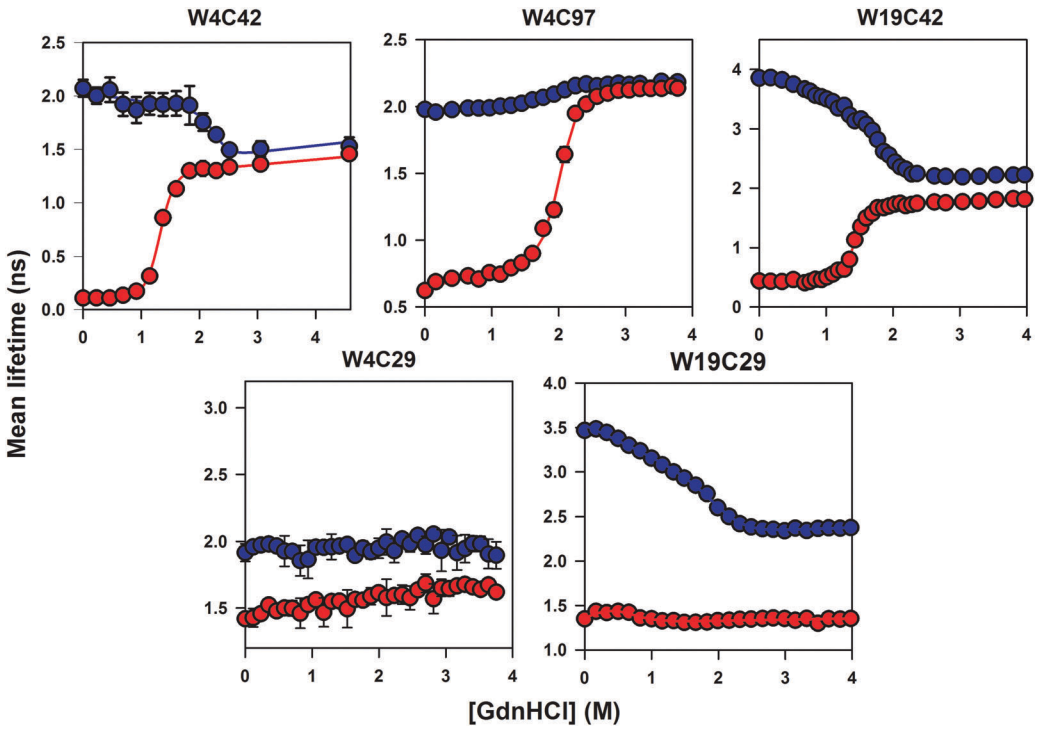

Fig. 4 Equilibrium unfolding transitions of MNEI monitored by time-resolved FRET. The fluorescence decay was monitored at 360 nm, using an excitation wavelength of 295 nm. Mean lifetimes were obtained from discrete analysis of the time-resolved data. Blue and red circles correspond to the data obtained for the unlabeled and TNB-labeled proteins, respectively. The solid line passing through each dataset is a non-linear, least-squares fit to a two-state unfolding model.[69] The error bars represent the standard deviations in the measurements of the same sample from three different acquisitions.

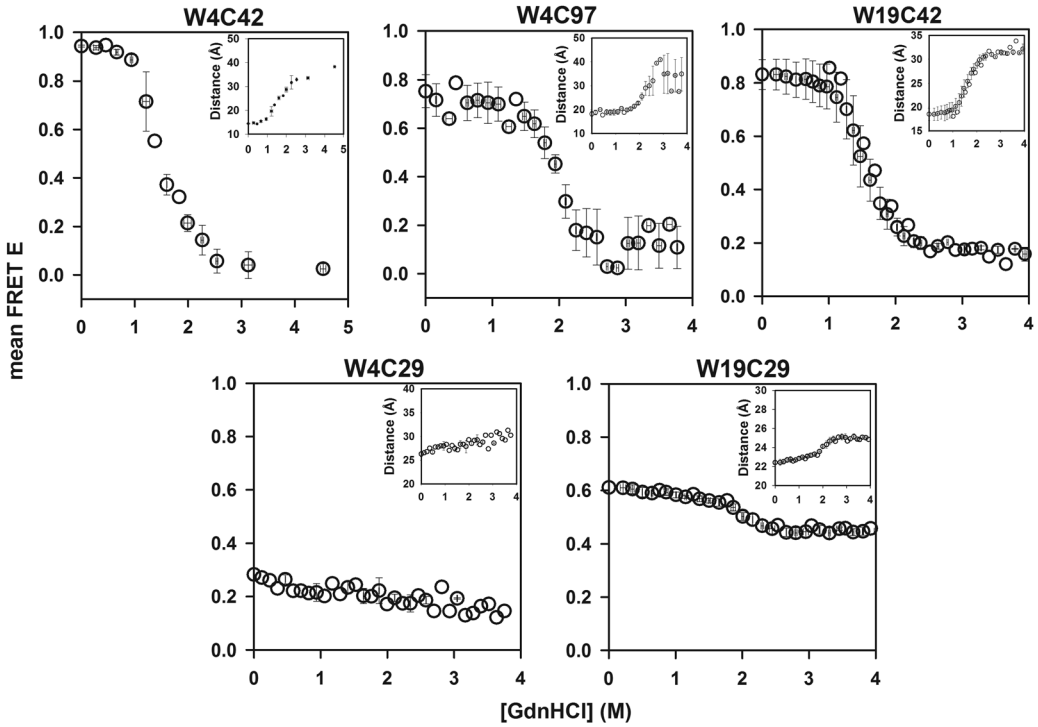

Fig. 5 Equilibrium unfolding transitions of MNEI monitored by time-resolved FRET. The fluorescence decay was monitored at 360 nm, using an excitation wavelength of 295 nm. Mean FRET efficiency is calculated using mean lifetimes that were obtained from discrete analysis of the time-resolved data for both unlabeled and TNB-labeled protein variants. The inset in each panel shows the dependence of the donor–acceptor distances on GdnHCl concentration. The distances were determined using measured R_0 values (Table S3, ESI,† eqn (9)) as described in the Materials and methods section. The error bars represent the standard error of the mean in the measurements of two independent samples.

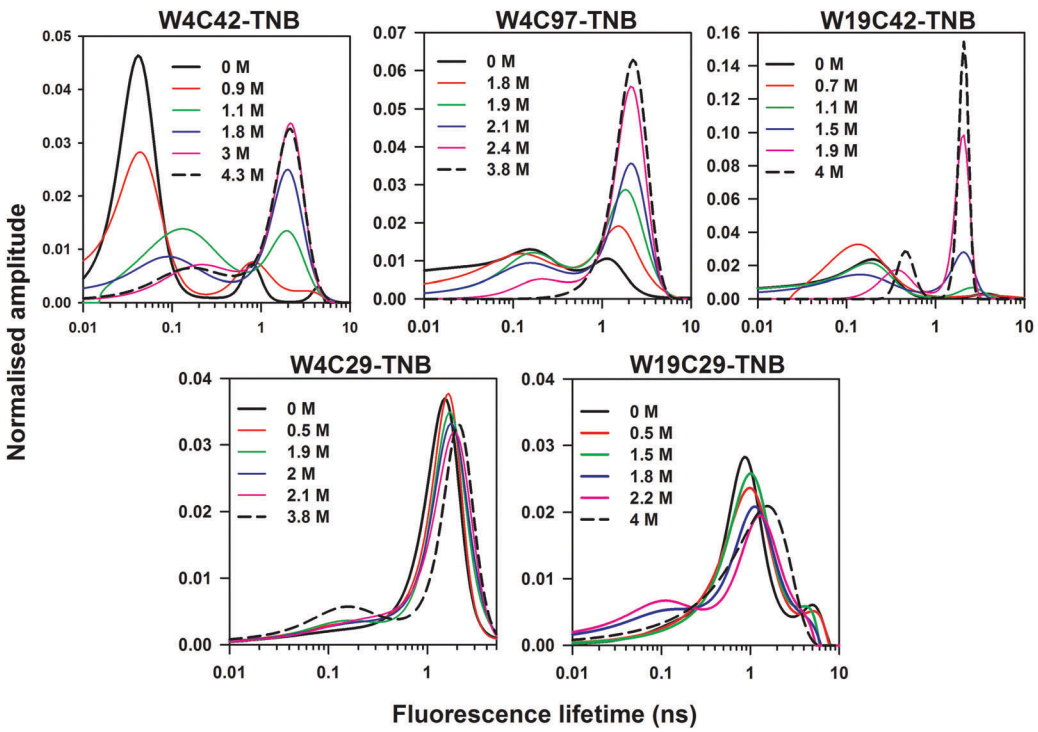

Fig. 6 MEM-derived fluorescence lifetime distributions of the TNB-labeled mutant variants of MNEI at different GdnHCl concentrations. Different colors correspond to various concentrations of GdnHCl, as described separately in each panel. The *x*-axis has been plotted on a log scale. The amplitude has been normalized to the sum of amplitudes for each distribution.

with increasing GdnHCl concentration was observed; (b) the peak position of the lifetime distributions of the U-like forms gradually shifted towards longer lifetimes for W4C42 and W4C97, but towards shorter lifetimes for W19C42; (c) the extent of the shift in the peak position with increasing GdnHCl concentration within each of the N-like or U-like populations varied from one TNB-labeled protein to another (Fig. 6 and Fig. S6, ESI†); and (d) the N-like and U-like ensembles inter-converted in a cooperative two-state manner, as indicated by a decrease in the population of the N-like forms at the expense of an increase in the population of U-like forms, with an increase in GdnHCl concentration (Fig. 6 and Fig. S5, ESI†). In contrast, the fluorescence lifetime distribution obtained for W4C29-TNB and W19C29-TNB, remained unimodal throughout the unfolding process (Fig. 6). The peak position of the lifetime distribution shifted gradually towards longer lifetimes with increasing GdnHCl concentration (Fig. 6 and Fig. S6, ESI†).

The distributions for all the unlabeled proteins remained unimodal during the process of unfolding (Fig. S4, ESI†). The lifetime distributions of unlabeled W4C42, W4C97 and W4C29 did not show a significant change upon unfolding, *i.e.*, the N and U states had very similar distributions. The unfolding was accompanied by a continuous increase, albeit very small, in the fluorescence lifetime as evident from the distribution moving towards a longer lifetime (Fig. S4 and S7, ESI†). For unlabeled W19C42 and W19C29, the distribution was found to move towards a shorter lifetime, in a continuous manner. This appeared to indicate that a specific change must occur in the environment in the vicinity of W19, which requires detailed investigation in the future.

It is important to note that to interpret changes in the fluorescence lifetime distributions obtained for the TNB-labeled protein variants, as changes in distance, it could be deemed necessary to properly disentangle the contribution from the respective unlabeled variants. However, in this study, the focus is on the differences seen between different TNB-labeled variants having very similar distributions for the unlabeled proteins. For example, W4C42, W4C97 and W4C29 show very similar lifetime distributions, as well as dependences of the lifetime distributions on [GdnHCl] (Fig. S4 and S7, ESI†). This suggests that the observation of differential cooperativity in the TNB-labeled variants (Fig. 6) must arise from differences in the nature of the structural segments monitored. Hence, there is no need, in the present study, to disentangle the fluorescence lifetime distributions of the unlabeled variant. A similar conclusion can be reached for W19C42 and W19C29 variants as well.

However, it is important to account for movement in the peak positions of the MEM distributions of the unlabeled variants before interpreting changes in the peak positions of TNB-labeled variants as changes in the dimensions of the proteins (Fig. 7). Thus, changes in the MEM peak distances have been quantified using the peak FRET efficiency values calculated by combining

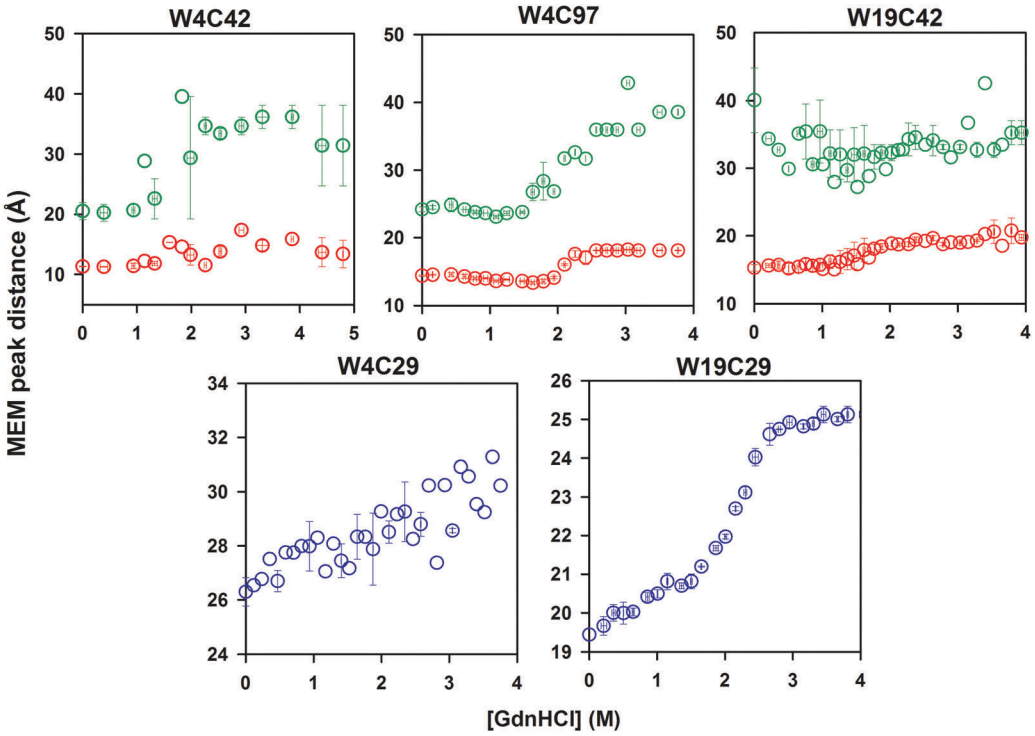

Fig. 7 Quantification of the expansion of the N state and U state ensemble. The movement in the MEM peak position for the TNB-labeled proteins and respective unlabeled proteins has been used to calculate the MEM peak FRET efficiency, which is transformed into distances using measured R_0 values (Table S3, ESI,† eqn (9)). Red and green circles represent the distances corresponding to the peak of the N-like (short lifetime) and U-like (long lifetime) species, respectively. The blue symbols in the panels for W4C29 and W19C29 correspond to the peak distances of the observed unimodal fluorescence lifetime distributions. The error bars represent the standard error of the mean of the measurements on two independent samples.

the changes in the MEM peak position/lifetime of both the unlabeled (Fig. S7, ESI†) and the TNB-labeled variants (Fig. S6, ESI†).

A two state N ↔ U model was inadequate for describing equilibrium unfolding

Assuming a two-state transition, the MEM distribution of the fluorescence lifetimes determined at all GdnHCl concentrations were fit using eqn (6), to the weighted sum of the mean of the MEM distributions determined for the N and U states (Fig. 8 and 9).

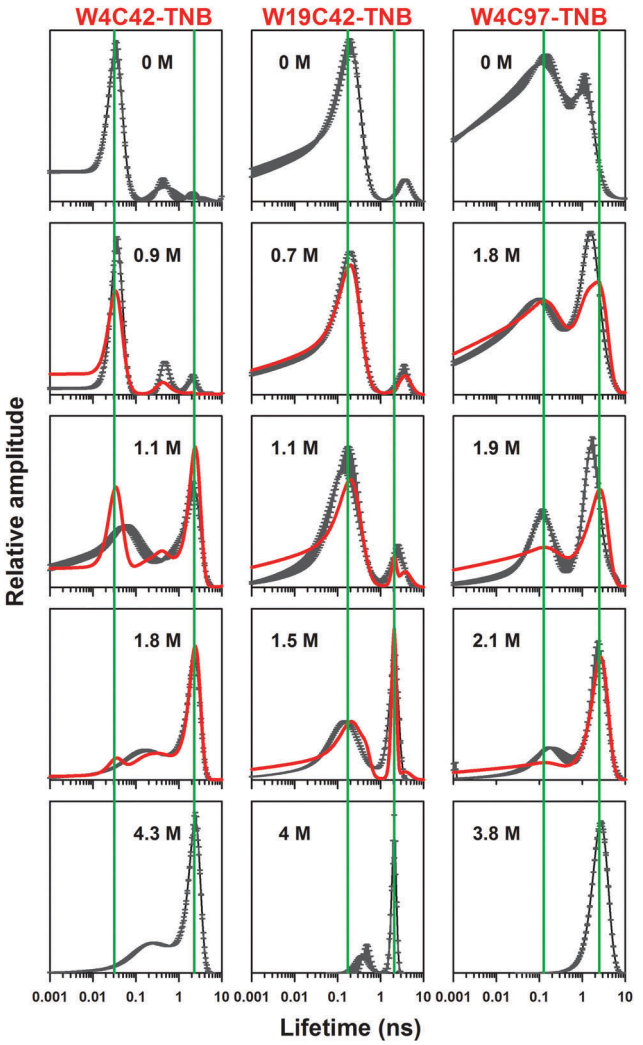

Fig. 8 Non-linear, least-squares fits to a two-state N ↔ U model for FRET pairs spanning different regions of the β-sheet. The MEM-derived fluorescence lifetime distributions obtained at different GdnHCl concentrations (indicated at the top of each panel) were fit to the linear sum of the native state N(τ) and unfolded state U(τ) fluorescence lifetime distributions (eqn (6)). The black distributions correspond to the experimentally determined distributions, and the red distributions are fits to the two-state model. The error bars are the standard deviations of the MEM distributions obtained from multiple data acquisitions on the same sample. The topmost and bottom-most panels in each column show the fluorescence lifetime distributions used as the native N(τ) and unfolded U(τ) protein basis spectra, respectively. The vertical green lines indicate the peak positions of the N-state and U-state lifetime distributions. The *x*-axis has been plotted on a log scale, and the *y*-axis units are arbitrary.

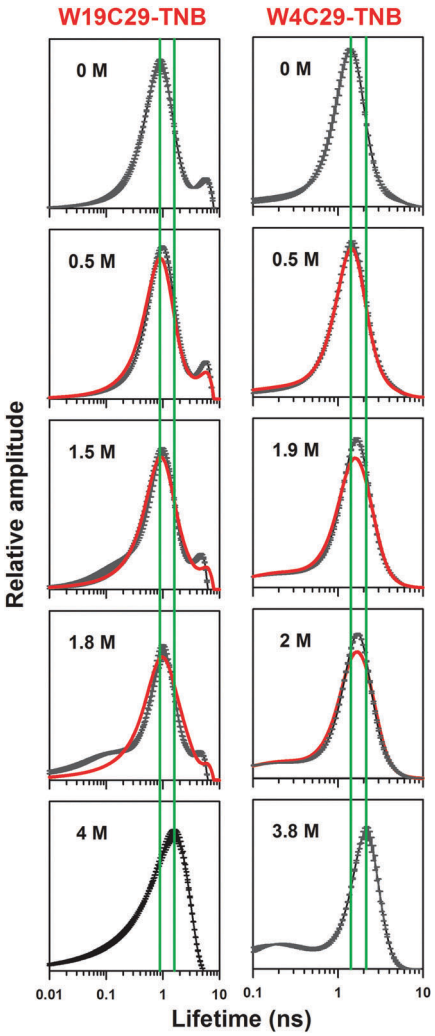

Fig. 9 Non-linear, least-squares fits to a two-state N ↔ U model for FRET pairs spanning different regions of the α-helix. The MEM-derived fluorescence lifetime distributions obtained at different GdnHCl concentrations (indicated at the top of each panel) were fit to the linear sum of the native state N(τ) and unfolded state U(τ) fluorescence lifetime distributions (eqn (6)). The black distributions correspond to the experimentally determined distributions, and the red distributions are fits to the two-state model. The error bars are the standard deviations of the MEM distributions obtained from multiple data acquisitions on the same sample. The topmost and bottom-most panels in each column show the fluorescence lifetime distributions used as the native N(τ) and unfolded U(τ) protein basis spectra, respectively. The vertical green lines indicate the peak positions of the N-state and U-state lifetime distributions. The *x*-axis has been plotted on a log scale, and the *y*-axis units are arbitrary.

Lifetime distributions obtained for GdnHCl concentrations >2.5 M, corresponding to the unfolded protein baseline region of the equilibrium unfolding curves, could be fitted to a linear combination of the native and unfolded state lifetime distributions in a reasonably acceptable manner as evident from the visual comparison of the simulated and experimental data (Fig. 8 and 9). The distributions of fluorescence lifetimes measured in the folding transition zones (corresponding to 0.5 to 2.5 M GdnHCl) of the equilibrium unfolding curve for most of the TNB-labeled proteins, could not, however, be fitted to a linear

combination of the distributions obtained for the N and U states, as the differences between the experimental and fitted data were beyond the experimental errors (Fig. 8). The observed significant deviations of the fitted lifetime distributions from the experimentally obtained distributions (Fig. 8, 9 and Fig. S10, ESI†) suggested that a two-state model was inadequate for describing the equilibrium unfolding transitions for at least three of the six distances monitored, in W4C42-TNB, W4C97-TNB and W19C42-TNB. The fit to a two-state model appeared to be relatively better for W4C29-TNB and W19C29-TNB as reflected in the rmsd values (Fig. S10, ESI†); however, for both W4C29-TNB and W19C29-TNB, an iso-lifetime point through which all the fluorescence lifetime distributions passed through, which is expected for a two-state unfolding transition, was absent (Fig. S11, ESI†). Also, the residuals for W4C29-TNB were not randomly distributed as expected for a good fit (Fig. S10, ESI†). It is to be noted here that the extent of deviation from a two-state model varied across different parts of the protein (rmsd values, Fig. S10, ESI†). It is important to note that these observations were consistent across two independent experiments, *i.e.* two different samples at each GdnHCl concentration (Fig. S8 and S9, ESI†), attesting to the robustness of the MEM analysis used here.

Apparent cooperative structural changes seen for some segments of the protein

In a bimodal MEM distribution, the sum of amplitudes for a particular peak, normalized to the sum of amplitudes for the full distribution, represents the fraction of the molecules that give rise to that peak. For W4C42-TNB, W19C42-TNB and W4C97-TNB, the sum of amplitudes under the distribution of N-like forms decreased at the expense of an increase in the sum of amplitudes under the distribution of U-like forms (Fig. 6 and Fig. S5, ESI†). For all three FRET pairs, a clear sigmoidal dependence of the relative populations corresponding to the N and U states on GdnHCl concentration was observed (Fig. S5, ESI†).

The transitions observed here were in good agreement with the major unfolding transitions for these proteins. It is to be noted that the separation of the MEM distributions of these three proteins into one component arising from N-like forms, and the other component arising from U-like forms, was validated by the observation that the resultant fraction U or fraction N plots (Fig. S5, ESI†) recapitulated the equilibrium unfolding transitions (Fig. 3 and 4 and Fig. S2, ESI†).

Continuous movement in the MEM-derived fluorescence lifetime distributions

Movement of the MEM fluorescence lifetime distribution peak positions with a change in GdnHCl concentration could be seen for all the TNB-labeled proteins (Fig. 6 and Fig. S6, ESI†). The changes in the peak position were evident for the unimodal distributions of W4C29-TNB and W19C29-TNB; and for both the N-like and U-like distributions of W4C42-TNB, W4C97-TNB and W19C42-TNB. The movement of the peak position with a change in GdnHCl concentration was very small for unlabeled W4C42, W4C97 and W4C29, but significant for W19C42 and W19C29 (Fig. S7, ESI†).

The dependence of the MEM peak position on GdnHCl concentration appeared to be sigmoidal in nature. A similar observation was made for the dependence of the discrete lifetime components on GdnHCl concentration (data not shown). As MEM lifetime distribution peaks are expected to correspond with the discrete components of fluorescence lifetime decay, it was not surprising that the movement in the MEM distribution peak had a sigmoidal dependence on GdnHCl concentration. The observation that the MEM peak lifetimes correspond well to the discrete components, justifies the use of the MEM peak positions for interpreting changes in the population distribution, and most likely structural changes, despite the asymmetry seen in the MEM distributions. However, the transitions observed with MEM peak movement for both the unlabeled (Fig. S7, ESI†) and TNB-labeled (Fig. S6, ESI†) proteins were different from the major unfolding transitions (Fig. 3, 4 and Fig. S2, ESI†), suggesting non-cooperativity in the unfolding process.

Discussion

Ensemble-averaging probes cannot usually distinguish between the different N-like forms that constitute the N state ensemble, and between the different U-like forms that constitute the U state ensemble. Nevertheless, they very often provide clues pointing to the heterogeneity inherent in the folding/unfolding transitions, such as significant dependences of the N state signal and U state signal on denaturant concentration. For example, in the case of MNEI, the fluorescence of the N state shows a strong dependence on GdnHCl concentration.[12] Moreover, the non-overlap of the CD-monitored and fluorescence-monitored unfolding transitions, seen for two variants of MNEI, W19C42-TNB and W19C29-TNB, suggests reduced cooperativity[5] in the equilibrium unfolding of MNEI (Fig. S3, ESI†).

To reveal the heterogeneity of folding/unfolding reactions, it becomes necessary to use a probe such as tr-FRET coupled with MEM analysis,[64,70] which can distinguish between the multitude of forms likely to be present at equilibrium, on the basis of differences in intra-molecular distances.[20,24] Since many forms may differ in only specific regions and not in all segments of structure, it becomes necessary to measure intra-molecular distances in multiple segments of the protein structure. In this study, five intra-molecular distances spanning different structural regions of the protein, were monitored.

The lack of cooperativity in the unfolding transition of MNEI is evident even in equilibrium studies

The observation (Fig. 3) that the equilibrium fluorescence monitored unfolding transitions of the five unlabeled protein variants and of the corresponding five labeled proteins are sigmoidal in nature suggested that unfolding of MNEI might be cooperative. Indeed, steady-state FRET (ss-FRET) measurements of the expansion of three intra-molecular distances also suggested cooperative unfolding. However, for two of the intra-molecular distances (W4C29 and W19C29), ss-FRET-monitored unfolding was not cooperative, but appeared to be gradual in nature

(insets, Fig. 3). The lack of cooperativity and presence of heterogeneity became more apparent in tr-FRET measurements of unfolding, when coupled to MEM analysis. Interestingly, for the segment monitored in W19C29, which showed cooperative unfolding (in terms of FRET efficiency and distance dependence) when probed by discrete analysis of tr-FRET (Fig. 5), a gradual unfolding transition was revealed upon population level MEM analysis (Fig. 6). The observation that for at least three of the five distances in W4C42-TNB, W4C97-TNB and W19C42-TNB, the MEM distribution of fluorescence lifetimes at intermediate GdnHCl concentrations could not be fit to the weighted sum of the native and unfolded protein MEM distributions (Fig. 8 and Fig. S10, ESI†) is strong evidence that multiple segments of the protein structure, each probed by an intra-molecular distance spanning it, unfold in a non-cooperative manner, albeit to different extents. The unfolding reaction did not appear to have limited cooperativity in that it could be described by the invocation of just one or a few intermediates; instead it appeared continuous in nature. In particular, the distances monitored in W4C29 and W19C29 clearly indicated that the segment spanning the helix swelled in a gradual manner (Fig. 5–7). Overall, it appears that different structural segments of the protein exhibit different degrees of cooperativity in their unfolding reactions, with two segments encompassing the helix displaying no cooperativity at all. These observations are consistent with previous kinetic studies of the unfolding of MNEI at high GdnHCl concentration, which had indicated that the protein unfolds by continuous swelling.[22] It is remarkable that such gradual swelling can also be revealed in an equilibrium study, as shown here. Differential segmental cooperativity had also been seen in the multi-site, tr-FRET monitored studies of the refolding kinetics of the core domain of the *E.coli* adenylate kinase.[51] In this study,[51] differential cooperativity was, however, not seen across different secondary structural units of the protein.

The helix in MNEI undergoes gradual structural change during unfolding

The observation that the two distances spanning the helix, in the mutant variants W4C29-TNB and W19C29-TNB, change in a complex and non-cooperative manner (Fig. 5–7) indicates that the helix dissolves in a gradual manner. However, it was possible that the continuous shift in the fluorescence lifetime distribution of the TNB-labeled protein, with increasing GdnHCl concentration, was indicative of a rapid exchange between the N state and U state, instead of continuous structural change accompanying the unfolding process. However, this is unlikely here, as (a) some of the fluorescence lifetime distributions near the midpoint of unfolding transitions could not be adequately described as a linear combination of the N and U state distributions (Fig. 9 and Fig. S9–S11, ESI†), (b) for fast exchange to occur, the rate of exchange would have to be faster than 10^{10} s^{-1}, the fastest rate of fluorescence decay, whereas conformational changes observed for proteins are found typically to be slower than 10^{9} s^{-1},[71] and (c) for W19C29-TNB, the apparent increase in the width of the distribution with increasing GdnHCl concentration (Fig. 6), is inconsistent with the fast exchange model. For W4C29-TNB, the complexity of the unfolding transition is evident even when monitored by ensemble-averaging probes such as ss-FRET (Fig. 3) and tr-FRET (Fig. 4 and 5). However, for W19C29-TNB, the heterogeneity in the unfolding gets masked when probed by discrete analysis of tr-FRET (mean lifetime, Fig. 5) measurements, highlighting the importance of high resolution methods (such as MEM) which provide population level information to resolve the heterogeneity inherent in protein folding–unfolding transitions.

Previous kinetic native state HX experiments[4] had suggested that a helical structure dissolves gradually but in three exponential phases, with the kinetic pauses separating the phases arising because of the need for specific packing interactions between the helix and the β-sheet to break.[4,19,25] The results of this study confirm that the helix unfolds in a completely gradual manner under equilibrium conditions. The gradual nature of the unfolding of the helix, when part of the protein, contrasts with the apparently cooperative nature of the unfolding of a helix, formed by an isolated peptide. For natural peptides in solution, the equilibrium unfolding reaction appears to be two-state in nature.[72,73] However, the kinetics of folding of helical peptides has been shown to follow a nucleation–elongation mechanism[74,75] where helix formation occurs in two steps; the formation of a helical nucleus is followed by the propagation of the helical structure. Earlier, it was shown that the formation of a helix becomes gradual when initiated by a permanent helical nucleus,[74] indicating that the process had appeared cooperative because the nucleation step involved crossing of a high free energy barrier from the U state. But it has also been suggested that the rate-limiting step in the folding of an α-helix is the escape from heterogeneous traps present in the U state rather than the nucleation process. Thus, when the heterogeneity in the U state is reduced by introducing structural constraints, a non-cooperative transition was seen.[42] This result brings out the importance of tertiary interactions in maintaining the integrity of the helical structure in a protein. It is also important to resolve which of the interactions, local or non-local, are responsible for giving rise to folding cooperativity in a protein. For an α-helix, the structure is stabilized mainly by local interactions/contacts, which need not form in a cooperative manner.[40,41,76]

β-sheet: the N and U state ensembles are separated by a free energy barrier

For the distances spanning the β-sheet, a clear cooperative conversion of the N-like forms to the U-like forms was observed (Fig. 6 and Fig. S5, ESI†). This suggested that there exists a significant free energy barrier between the N state ensemble and the U state ensemble. The observation that the relative populations of the N-like and U-like ensembles depend upon the GdnHCl concentration is similar to previous observation made on the unfolding of adenylate kinase, where sm-FRET measurements showed that the amounts of different cooperatively exchanging intermediates could be tuned by varying the denaturant concentration.[77]

Similar levels of segmental cooperativity had been observed previously in equilibrium studies of the unfolding of barstar[35]

and the SH3 domain of PI3K,[24] and also in a kinetic study of the unfolding of MNEI,[22] exclusively when monitored by FRET pairs spanning the β-sheet. A kinetic HX study has shown previously that the unfolding of MNEI occurs in a gradual manner in the absence of denaturant, and that the reaction gains cooperativity in the presence of even small amounts of denaturants.[25] It was shown that two of the β-strands (β2 & β3) unfold cooperatively in the presence of denaturant, while the rest of the structure still unfolds gradually.[4]

The structural basis of this cooperativity is not really understood. The cooperativity might be arising due to the barrier associated with the formation of non-local interactions that bring different strands of a β-sheet together, as the amino acid residues which participate in inter-β-strand interactions are distant from each other along the primary sequence of the protein.[76] A mutagenesis approach that perturbs such long-range contacts might help in understanding the structural basis of cooperativity exclusive to the β-sheet of MNEI and other small proteins.

The N state ensemble swells and becomes more heterogeneous with an increase in GdnHCl concentration

For W4C42-TNB, W19C42-TNB and W4C97-TNB, the distances corresponding to the peaks of the fluorescence lifetime distributions of the N state ensemble shift gradually to higher lifetime values when the GdnHCl concentration is increased from 0 to 2 M (Fig. 6 and 7). This observation suggests that the N state ensemble swells gradually in this range of GdnHCl concentration (Fig. 7). This is unlikely to be a general solvent effect as (a) such a drastic change (fractional increase) is not observed for the unlabeled proteins (thus reflected as a change in FRET efficiency and distance); (b) the extent of change is different for different segments monitored (Fig. 7), and (c) the widths of the fluorescence lifetime distributions increase with increasing GdnHCl concentrations (Fig. 6, 8 and Fig. S8, ESI†). The N state appears to swell to a more heterogeneous molten globule-like form. Early theoretical studies of equilibrium unfolding[78,79] had suggested that the initial step in unfolding is swelling of the native state, which disrupts the packing interactions that maintain the integrity of the protein structure. A kinetic study[56] had supported this theory that unfolding begins by formation of a dry molten globule. A previous tr-FRET study of the equilibrium unfolding of barstar had shown that the N state continues to swell until the protein attains the dimensions of a wet molten globule.[35] This suggests that non-cooperative structural changes with very little energy cost, may be a prerequisite to the major cooperative unfolding transition. An interesting observation is that all segments in the native protein swell differentially (Fig. 7). In future, it will be important to identify the interactions that retard segmental swelling in some regions of the native protein structure. It is to be noted that no significant change in the dimensions of the N state ensemble could be observed in the case of W4C42. This is likely because the FRET pair (W-TNB) used, with an R_0 value determined to be 23.1 Å, is not sensitive enough to monitor changes in the distance (10–14 Å) separating donor and acceptor in W4C42-TNB.

The U state ensemble expands differentially in different sequence segments

The gradual shift in the fluorescence lifetime distributions arising from the U state ensemble for W4C42-TNB, W4C97-TNB and W19C42-TNB (Fig. 6 and Fig. S6, ESI†) indicates expansion of the U state, with increasing GdnHCl concentration. For W19C42-TNB, the lifetime distribution moved, however, towards shorter lifetimes with increasing GdnHCl concentration. This decrease in fluorescence lifetime could have been interpreted as compaction, except that a similar shift occurred for the fluorescence lifetime distributions of the unlabeled protein (Fig. S4 and S7, ESI†). When the distances were calculated from the peak lifetimes (Fig. 7), it, however, became evident that the distribution spanning the W19 to C42-TNB in the U state does not undergo any major change in its dimensions with increasing GdnHCl concentration. The observation that different intra-molecular distances expand over different ranges of GdnHCl concentration (Fig. 7) suggests that the different segments of the sequence might encompass residual structures of varying stability, and that melting of the residual structure contributes to the expansion of the U state ensemble. In the case of the unfolded state of fyn-SH3, sm-FRET measurements of five intra-molecular distances had suggested the presence of residual structure in the U state.[80] Gradual expansion of the U state ensemble with an increase in denaturant concentration has been seen for many proteins.[22,35,52,81–86]

Heterogeneity in the unfolding reaction

The use of specific intra-molecular distances to probe the equilibrium unfolding of specific segments of the structure of MNEI, has brought out the heterogeneity of an unfolding transition that otherwise appeared two-state when probed by traditional ensemble-averaging probes. In three segments of the protein; monitored by distances spanning W4 to C42-TNB, W4 to C97-TNB and W19 to C42-TNB, unfolding has characteristics of both gradual and all-or-none change. These segments in the N state ensemble swell upon addition of GdnHCl, suggesting that unfolding commences by an apparently higher-order transition through a multitude of N-like forms. It appears that a first-order transition to the U state ensemble occurs only after these segments have swollen sufficiently for many of the packing, and possibly long range, interactions that maintain the integrity of the N state structure, to have broken. In the U state ensemble, the more expanded conformations become favored as the GdnHCl concentration is increased. It is perhaps not surprising that the segments showing such heterogeneous unfolding behavior include the β-sheet structure and the interface between the helix and β-sheet. Packing interactions between side chains of β-strands are known to be important for stability,[46] and are known to play an important role in the unfolding of MNEI.[4,19,25] Similar heterogeneity in the segments reported on by the W4 to C42-TNB distance, and by the W4 to C97-TNB distance, had earlier been seen in kinetic studies of the unfolding of MNEI.[22] A similar non-uniform expansion of different intra-molecular distances had been seen earlier as well for the U state of barstar[81,82] and acyl-CoA-binding protein.[52]

It should be noted that from only an equilibrium study, it is difficult to deduce any quantitative model to describe a complex unfolding reaction, because a quantitative model requires the determination of the temporal sequence of structural changes, which is only possible from a kinetic study. Nevertheless, the qualitative model proposed here in this equilibrium study of the unfolding of MNEI, is consistent with the quantitative model described earlier based on a kinetic study of the unfolding of MNEI.[22] In fact, the complexity revealed in the current equilibrium study by the use of FRET pairs not used in the earlier kinetic study, suggests that the quantitative model will need to be expanded; for that, additional kinetic studies will need to be carried out in future.

In the two segments that encompass the helix, and which are reported on by the distributions of the distances separating C29-TNB from W4 and W19, there is no evidence for the helix unfolding using a barrier-limited process. Instead unfolding appears to be completely gradual. It would appear that as the packing at the helix-β-sheet interface is broken upon swelling of the β-sheet, the helix is stabilized *via* local interactions which break in a gradual manner. It should be noted that although the changes in FRET efficiency and distances that monitor the unfolding of the helix are small, the changes are robust and reproducible.

It is important to note that the degree of heterogeneity mapped here would not be measurable by any other probe because high structural sensitivity (in the range of 10 to 30 Å) and high temporal (ps–ns) resolution were required to dissect out the differences in segment-specific cooperativity. sm-FRET measurements would be unable to reveal the heterogeneity inherent in this reaction as (a) the structural sensitivity in sm-FRET measurements is restricted to distances in the range of 30 to 70 Å because of the fluorophores that have to be used,[87] and (b) a fast exchanging species can get masked due to the poor temporal resolution (100 μs) of sm-FRET measurements.[88]

Conclusions

The present study focused on understanding the thermodynamic cooperativity of the unfolding of MNEI. Earlier equilibrium studies had failed to reveal the complexity inherent in the unfolding of MNEI, which therefore appeared to be describable as two state in nature. This study not only falsifies the two state model, but also provides a region-specific measure of the extent of deviation from the two-state model. Unfolding across five different distances was monitored at equilibrium, using tr-FRET. It was found that a multitude of equilibrium intermediates was populated during the denaturant-induced unfolding of MNEI. It was also observed that both the N state and the U state expanded gradually with increasing denaturant concentration. Interestingly, the sole α-helix present in the protein appeared to unfold in a completely non-cooperative manner which was quite distinct from the rest of the structural regions.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank MK Mathew and the members of our laboratory for discussions. We thank Rama Reddy Goluguri and Madhuri for technical help with experiments. JBU is a recipient of a JC Bose National 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

1 K. A. Dill, K. M. Fiebig and H. S. Chan, Cooperativity in protein-folding kinetics, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 1942–1946.
2 C. R. Babu, V. J. Hilser and A. J. Wand, Direct access to the cooperative substructure of proteins and the protein ensemble *via* cold denaturation, *Nat. Struct. Mol. Biol.*, 2004, **11**, 352–357.
3 V. Muñoz, L. A. Campos and M. Sadqi, Limited cooperativity in protein folding, *Curr. Opin. Struct. Biol.*, 2016, **36**, 58–66.
4 P. Malhotra and J. B. Udgaonkar, Secondary Structural Change Can Occur Diffusely and Not Modularly during Protein Folding and Unfolding Reactions, *J. Am. Chem. Soc.*, 2016, **138**, 5866–5878.
5 P. Malhotra and J. B. Udgaonkar, How cooperative are protein folding and unfolding transitions?, *Protein Sci.*, 2016, 1–18.
6 P. S. Kim and R. L. Baldwin, Specific intermediates in the folding recations of small proteins and the mechanism of protein folding pathways, *Annu. Rev. Biochem.*, 1982, **51**, 459–489.
7 J. B. Udgaonkar and R. L. Baldwin, NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A, *Nature*, 1988, **335**, 694–699.
8 H. Roder, G. A. Elöve and S. W. Englander, Structural characterization of folding intermediates in cytochrome *c* by H-exchange labelling and proton NMR, *Nature*, 1988, **335**, 700–704.
9 M. C. Shastry and J. B. Udgaonkar, The folding mechanism of barstar: evidence for multiple pathways and multiple intermediates, *J. Mol. Biol.*, 1995, **247**, 1013–1027.
10 F. N. Zaidi, U. Nath and J. B. Udgaonkar, Multiple intermediates and transition states during protein unfolding, *Nat. Struct. Biol.*, 1997, **4**, 1016–1024.
11 I. E. Sánchez and T. Kiefhaber, Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding, *J. Mol. Biol.*, 2003, **325**, 367–376.
12 A. K. Patra and J. B. Udgaonkar, Characterization of the folding and unfolding reactions of single-chain monellin: Evidence for multiple intermediates and competing pathways, *Biochemistry*, 2007, **46**, 11727–11743.
13 J. B. Udgaonkar, Multiple routes and structural heterogeneity in protein folding, *Annu. Rev. Biophys.*, 2008, **37**, 489–510.
14 S. S. Sarkar, J. B. Udgaonkar and G. Krishnamoorthy, Unfolding of a small protein proceeds via dry and wet

globules and a solvated transition state, *Biophys. J.*, 2013, **105**, 2392–2402.
15 R. R. Goluguri and J. B. Udgaonkar, Microsecond Rearrangements of Hydrophobic Clusters in an Initially Collapsed Globule Prime Structure Formation during the Folding of a Small Protein, *J. Mol. Biol.*, 2016, **428**, 3102–3117.
16 S. E. Jackson and A. R. Fersht, Folding of chymotrypsin inhibitor 2. 2. Influence of proline isomerization on the folding kinetics and thermodynamic characterization of the transition state of folding, *Biochemistry*, 1991, **30**, 10436–10443.
17 K. W. Plaxco, K. T. Simons, I. Ruczinski and D. Baker, Topology, stability, sequence, and length: Defining the determinants of two-state protein folding kinetics, *Biochemistry*, 2000, **39**, 11177–11183.
18 H. S. Chan and K. A. Dill, Protein folding in the landscape perspective: Chevron plots and non-rrhenius kinetics, *Proteins: Struct., Funct., Genet.*, 1998, **30**, 2–33.
19 P. Malhotra and J. B. Udgaonkar, High-energy intermediates in protein unfolding characterized by thiol labeling under nativelike conditions, *Biochemistry*, 2014, **53**, 3608–3620.
20 G. S. Lakshmikanth, K. Sridevi, G. Krishnamoorthy and J. B. Udgaonkar, Structure is lost incrementally during the unfolding of barstar, *Nat. Struct. Biol.*, 2001, **8**, 799–804.
21 E. V. Kuzmenkina, C. D. Heyes and G. U. Nienhaus, Single-molecule FRET Study of Denaturant Induced Unfolding of RNase H, *J. Mol. Biol.*, 2006, **357**, 313–324.
22 S. K. Jha, D. Dhar, G. Krishnamoorthy and J. B. Udgaonkar, Continuous dissolution of structure during the unfolding of a small protein, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 11113–11118.
23 Z. Liu, G. Reddy, E. P. O'Brien and D. Thirumalai, Collapse kinetics and chevron plots from simulations of denaturant-dependent folding of globular proteins, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 7787–7792.
24 M. Kishore, G. Krishnamoorthy and J. B. Udgaonkar, Critical evaluation of the two-state model describing the equilibrium unfolding of the PI3K SH3 domain by time-resolved fluorescence resonance energy transfer, *Biochemistry*, 2013, **52**, 9482–9496.
25 P. Malhotra and J. B. Udgaonkar, Tuning Cooperativity on the Free Energy Landscape of Protein Folding, *Biochemistry*, 2015, **54**, 3431–3441.
26 A. Narayan, L. A. Campos, S. Bhatia, D. Fushman and A. N. Naganathan, Graded Structural Polymorphism in a Bacterial Thermosensor Protein, *J. Am. Chem. Soc.*, 2017, **139**, 792–802.
27 A. Holtzer, M. E. Holtzer, E. G. Lovett and D. A. d'Avignon, Thermal unfolding in a GCN4-like leucine zipper: ^{13}C alpha NMR chemical shifts and local unfolding curves, *Biophys. J.*, 1997, **73**, 1031–1041.
28 M. Sadqi, D. Fushman and V. Muñoz, Atom-by-atom analysis of global downhill protein folding, *Nature*, 2006, **442**, 317–321.
29 A. Chatterjee, P. M. Krishna Mohan, A. Prabhu, A. Ghosh-Roy and R. V. Hosur, Equilibrium unfolding of DLC8 monomer by urea and guanidine hydrochloride: Distinctive global and residue level features, *Biochimie*, 2007, **89**, 117–134.
30 O. Julien, S. Chatterjee, A. Thiessen, S. P. Graether and B. D. Sykes, Differential stability of the bovine prion protein upon urea unfolding, *Protein Sci.*, 2009, **18**, 2172–2182.
31 P. M. K. Mohan, S. Chakraborty and R. V. Hosur, NMR investigations on residue level unfolding thermodynamics in DLC8 dimer by temperature dependent native state hydrogen exchange, *J. Biomol. NMR*, 2009, **44**, 1–11.
32 L. Sborgi, A. Verma, S. Piana, K. Lindorff-Larsen, M. Cerminara, C. M. Santiveri, D. E. Shaw, E. De Alba and V. Muñoz, Interaction Networks in Protein Folding via Atomic-Resolution Experiments and Long-Time-Scale Molecular Dynamics Simulations, *J. Am. Chem. Soc.*, 2015, **137**, 6506–6516.
33 J. Sabelko, J. Ervin and M. Gruebele, Observation of strange kinetics in protein folding, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 6031–6036.
34 J. Song, N. Jamin, B. Gilquin, C. Vita and A. Ménez, A gradual disruption of tight side-chain packing: 2D 1H-NMR characterization of acid-induced unfolding of CHABII, *Nat. Struct. Biol.*, 1999, **6**, 129–134.
35 G. S. Lakshmikanth, K. Sridevi, G. Krishnamoorthy and J. B. Udgaonkar, Structure is lost incrementally during the unfolding of barstar, *Nat. Struct. Biol.*, 2001, **8**, 799–804.
36 J. M. Sanchez-ruiz and V. Mun, Experimental Identi cation of Downhill Protein Folding, *Science*, 2002, **298**, 2191–2195.
37 K. K. Sinha and J. B. Udgaonkar, Barrierless evolution of structure during the submillisecond refolding reaction of a small protein, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 7998–8003.
38 S. A. Waldauer, O. Bakajin, T. Ball, Y. Chen, S. J. Decamp, M. Kopka, M. Jäger, V. R. Singh, W. J. Wedemeyer, S. Weiss, S. Yao and L. J. Lapidus, Ruggedness in the folding landscape of protein L, *HFSP J.*, 2008, **2**, 388–395.
39 A. H. Wani and J. B. Udgaonkar, Native state dynamics drive the unfolding of the SH3 domain of PI3 kinase at high denaturant concentration, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 20711–20716.
40 T. R. Weikl and K. A. Dill, Transition-States in Protein Folding Kinetics: The Structural Interpretation of Φ-values, *J. Mol. Biol.*, 2007, **365**, 1578–1586.
41 A. N. Naganathan and V. Muñoz, Insights into protein folding mechanisms from large scale analysis of mutational effects, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 8611–8616.
42 J. a Ihalainen, B. Paoli, S. Muff, E. H. G. Backus, J. Bredenbeck, G. A. Woolley, A. Caflisch and P. Hamm, Alpha-Helix folding in the presence of structural constraints, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 9588–9593.
43 A. P. Capaldi and S. E. Radford, Kinetic Studies of b-sheet protein folding, *Curr. Opin. Struct. Biol.*, 1998, **8**, 86–92.
44 D. Samuel, T. K. S. Kumar, K. Balamurugan, W. Y. Lin, D. H. Chin and C. Yu, Structural Events during the Refolding of an All β-Sheet Protein, *J. Biol. Chem.*, 2001, **276**, 4134–4141.
45 S. Deechongkit, H. Nguyen, E. T. Powers, P. E. Dawson, M. Gruebele and J. W. Kelly, Context-dependent contributions

of backbone hydrogen bonding to beta-sheet folding energetics, *Nature*, 2004, **430**, 101–105.

46 V. M. Kung, G. Cornilescu and S. H. Gellman, Impact of Strand Number on Parallel β-Sheet Stability, *Angew. Chem., Int. Ed. Engl.*, 2015, **54**, 14336–14339.

47 N. Aghera and J. B. Udgaonkar, Stepwise Assembly of β-Sheet Structure during the Folding of an SH3 Domain Revealed by a Pulsed Hydrogen Exchange Mass Spectrometry Study, *Biochemistry*, 2017, **56**, 3754–3769.

48 X. Michalet, S. Weiss and M. Jäger, Single-molecule fluorescence studies of protein folding and conformational dynamics, *Chem. Rev.*, 2006, **106**, 1785–1813.

49 B. Schuler and H. Hofmann, Single-molecule spectroscopy of protein folding dynamics—expanding scope and timescales, *Curr. Opin. Struct. Biol.*, 2013, **23**, 36–47.

50 P. S. Eis and J. R. Lakowicz, Time-resolved energy transfer measurements of donor-acceptor distance distributions and intramolecular flexibility of a CCHH zinc finger peptide, *Biochemistry*, 1993, **32**, 7981–7993.

51 V. Ratner, D. Amir, E. Kahana and E. Haas, Fast collapse but slow formation of secondary structure elements in the refolding transition of *E. coli* adenylate kinase, *J. Mol. Biol.*, 2005, **352**, 683–699.

52 V. A. Voelz, M. Jäger, S. Yao, Y. Chen, L. Zhu, S. A. Waldauer, G. R. Bowman, M. Friedrichs, O. Bakajin, L. J. Lapidus, S. Weiss and V. S. Pande, Slow Unfolded-State Structuring in Acyl-CoA Binding Protein Folding Revealed by Simulation and Experiment, *J. Am. Chem. Soc.*, 2012, **134**, 12565–12577.

53 N. Aghera, N. Earanna and J. B. Udgaonkar, Equilibrium unfolding studies of monellin: The double-chain variant appears to be more stable than the single-chain variant, *Biochemistry*, 2011, **50**, 2434–2444.

54 N. Aghera and J. B. Udgaonkar, Kinetic Studies of the Folding of Heterodimeric Monellin: Evidence for Switching between Alternative Parallel Pathways, *J. Mol. Biol.*, 2012, **420**, 235–250.

55 T. Kimura, A. Maeda, S. Nishiguchi, K. Ishimori, I. Morishima, T. Konno, Y. Goto and S. Takahashi, Dehydration of main-chain amides in the final folding step of single-chain monellin revealed by time-resolved infrared spectroscopy, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 13391–13396.

56 S. K. Jha and J. B. Udgaonkar, Direct evidence for a dry molten globule intermediate during the unfolding of a small protein, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 12289–12294.

57 S. K. Jha, A. Dasgupta, P. Malhotra and J. B. Udgaonkar, Identification of multiple folding pathways of monellin using pulsed thiol labeling and mass spectrometry, *Biochemistry*, 2011, **50**, 3062–3074.

58 N. Aghera, I. Dasgupta and J. B. Udgaonkar, A buried ionizable residue destabilizes the native state and the transition state in the folding of monellin, *Biochemistry*, 2012, **51**, 9058–9066.

59 N. Aghera and J. B. Udgaonkar, The utilization of competing unfolding pathways of monellin is dictated by enthalpic barriers, *Biochemistry*, 2013, **52**, 5770–5779.

60 R. R. Goluguri and J. B. Udgaonkar, Rise of the Helix from a Collapsed Globule during the Folding of Monellin, *Biochemistry*, 2015, **54**, 5356–5365.

61 J. Skilling and R. Bryan, Maximum entropy image reconstruction: general algorithm, *Mon. Not. R. Astron. Soc.*, 1984, **211**, 111–124.

62 A. K. Livesey and J. C. Brochon, Analyzing the distribution of decay constants in pulse-fluorimetry using the maximum entropy method, *Biophys. J.*, 1987, **52**, 693–706.

63 J. C. Brochon, Maximum entropy method of data analysis in time-resolved spectroscopy, *Methods Enzymol.*, 1994, **240**, 262–311.

64 R. Swaminathan and N. Periasamy, 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.*, 1996, **108**, 39–49.

65 J. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, New York, 3rd edn, 2006, pp. 443–472.

66 K. Sridevi and J. B. Udgaonkar, Surface expansion is independent of and occurs faster than core solvation during the unfolding of barstar, *Biochemistry*, 2003, **42**, 1551–1563.

67 E. Haas, E. Katchalski-Katzir and I. Z. Steinberg, Effect of the orientation of donor and acceptor on the probability of energy transfer involving electronic transitions of mixed polarization, *Biochemistry*, 1978, **17**, 5064–5070.

68 S. Ramachandran and J. B. Udgaonkar, Stabilization of barstar by chemical modification of the buried cysteines, *Biochemistry*, 1996, **35**, 8776–8785.

69 V. R. Agashe and J. B. Udgaonkar, Thermodynamics of Denaturation of Barstar: Evidence for Cold Denaturation and Evaluation of the Interaction with Guanidine Hydrochloride, *Biochemistry*, 1995, **34**, 3286–3299.

70 W. M. Shih, Z. Gryczynski, J. R. Lakowicz and J. A. Spudich, A FRET-based sensor reveals large ATP hydrolysis-induced conformational changes and three distinct states of the molecular motor myosin, *Cell*, 2000, **102**, 683–694.

71 R. Ishima and D. Torchia, Protein dynamics from NMR, *Nat. Struct. Mol. Biol.*, 2000, **76**, 145–152.

72 I. K. Lednev, A. S. Karnoup, M. C. Sparrow and S. A. Asher, α-Helix peptide folding and unfolding activation barriers: A nanosecond UV resonance raman study, *J. Am. Chem. Soc.*, 1999, **121**, 8074–8086.

73 D. S. Talaga, W. L. Lau, H. Roder, J. Tang, Y. Jia, W. F. DeGrado and R. M. Hochstrasser, Dynamics and folding of single two-stranded coiled-coil peptides studied by fluorescent energy transfer confocal microscopy, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 13021–13026.

74 J. M. Scholtz and R. L. Baldwin, The Mechanism of alpha-Helix Formation by Peptides, *Annu. Rev. Biophys. Biomol. Struct.*, 1992, **21**, 95–118.

75 U. R. Doshi and V. Muñoz, The principles of α-helix formation: Explaining complex kinetics with nucleation-elongation theory, *J. Phys. Chem. B*, 2004, **108**, 8497–8506.

76 V. I. Abkevich, A. M. Gutin and E. I. Shakhnovich, Impact of local and non-local interactions on thermodynamics and kinetics of protein folding, *J. Mol. Biol.*, 1995, **252**, 460–471.

77 M. Pirchi, G. Ziv, I. Riven, S. S. Cohen, N. Zohar, Y. Barak and G. Haran, Single-molecule fluorescence spectroscopy maps the folding landscape of a large protein, *Nat. Commun.*, 2011, **2**, 493.

78 E. I. Shakhnovich and A. V. Finkelstein, Theory of cooperative transitions in protein molecules. I. Why denaturation of globular protein is a first-order phase transition, *Biopolymers*, 1989, **28**, 1667–1680.

79 A. V. Finkelstein and E. I. Shakhnovich, Theory of cooperative transitions in protein molecules. II. Phase diagram for a protein molecule in solution, *Biopolymers*, 1989, **28**, 1681–1694.

80 E. R. McCarney, J. H. Werner, S. L. Bernstein, I. Ruczinski, D. E. Makarov, P. M. Goodwin and K. W. Plaxco, Site-specific dimensions across a highly denatured protein; a single molecule study, *J. Mol. Biol.*, 2005, **352**, 672–682.

81 K. K. Sinha and J. B. Udgaonkar, 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.*, 2005, **353**, 704–718.

82 A. M. Saxena, J. B. Udgaonkar and G. Krishnamoorthy, Characterization of Intra-molecular Distances and Site-specific Dynamics in Chemically Unfolded Barstar: Evidence for Denaturant-dependent Non-random Structure, *J. Mol. Biol.*, 2006, **359**, 174–189.

83 S. Mukhopadhyay, R. Krishnan, E. A. Lemke, S. Lindquist and A. A. Deniz, A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 2649–2654.

84 A. C. M. Ferreon, C. R. Moran, Y. Gambin and A. A. Deniz, Single-molecule fluorescence studies of intrinsically disordered proteins, *Methods Enzymol.*, 2010, **472**, 179–204.

85 H. Hofmann, A. Soranno, A. Borgia, K. Gast, D. Nettels and B. Schuler, Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 16155–16160.

86 M. Aznauryan, L. Delgado, A. Soranno, D. Nettels, J. Huang, A. M. Labhardt, S. Grzesiek and B. Schuler, Comprehensive structural and dynamical view of an unfolded protein from the combination of single-molecule FRET, NMR, and SAXS, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, 201607193.

87 A. C. M. Ferreon and A. A. Deniz, Protein folding at single-molecule resolution, *Biochim. Biophys. Acta, Proteins Proteomics*, 2011, **1814**, 1021–1029.

88 B. Schuler, E. A. Lipman and W. A. Eaton, Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy, *Nature*, 2002, **419**, 743–747.

