Continuous dissolution of structure during the unfolding of a small protein

Santosh Kumar Jha[a], Deepak Dhar[b,1], Guruswamy Krishnamoorthy[c,1], and Jayant B. Udgaonkar[a,1]

[a]National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore 560065, India; and Departments of [b]Theoretical Physics and [c]Chemical Sciences, Tata Institute of Fundamental Research, Mumbai 400005, India

Edited by Ken A. Dill, University of California, San Francisco, CA, and approved May 1, 2009 (received for review December 11, 2008)

The unfolding kinetics of many small proteins appears to be first order, when measured by ensemble-averaging probes such as fluorescence and circular dichroism. For one such protein, monellin, it is shown here that hidden behind this deceptive simplicity is a complexity that becomes evident with the use of experimental probes that are able to discriminate between different conformations in an ensemble of structures. In this study, the unfolding of monellin has been probed by measurement of the changes in the distributions of 4 different intramolecular distances, using a multisite, time-resolved fluorescence resonance energy transfer methodology. During the course of unfolding, the protein molecules are seen to undergo slow and continuous, diffusive swelling. The swelling process can be modeled as the slow diffusive swelling of a Rouse-like chain with some additional noncovalent, intramolecular interactions. Here, we show that specific structure is lost during the swelling process gradually, and not in an all-or-none manner, during unfolding.

diffusive swelling | distance distribution | gradual unfolding | Rouse model | time-resolved FRET

It has been argued that the energy landscape encountered by a protein for traversing between the native (N) and unfolded (U) states, can be quite complex (1–6). However, many protein folding reactions, and most particularly, unfolding reactions appear simple and apparently "two-state," with the kinetics describable by transition state theory. According to transition state theory, the reactions are slowed down by a single dominant free energy barrier describable in terms of a single reaction coordinate. In this simple approach, an unfolding protein molecule does not spend appreciable time in conformations intermediate between N and U. However, there has been growing realization that this is not a good approximation in many cases, even for some relatively small proteins. In fact, the thermodynamic transition between the N and U states is expected to be a continuous transition, which is also rounded off due to finite size effects, implying that a sharp distinction between the N and U states is not possible (7, 8). In the alternative viewpoint, the free energy landscape is multidimensional, over which a large number of small ($<3\ k_BT$) barriers are distributed (9).

An earlier theoretical study had suggested that structural transitions within the molten globular states formed during unfolding, including globule swelling, occur in a gradual manner (10). Recent equilibrium unfolding studies using high-resolution structural probes or single-molecule fluorescence detection, have also indicated that unfolding may occur in several steps (11) and even gradually via a continuum of intermediate forms (12–17). Nevertheless, definitive kinetic evidence for a gradual protein unfolding or folding reaction has been scarce, and only indirect (18–20).

To distinguish experimentally between a two-state and a gradual unfolding reaction has been a difficult challenge, because of the problem in differentiating between, and quantifying the relative populations of different conformations present together at any time during unfolding. One solution to the problem is to carry out time-resolved fuorescence resonance energy transfer (TR-FRET) experiments (14, 21–25), coupled to the maximum entropy method (MEM) of analysis (14, 23, 24). In the TR-FRET technique, the time dependent decay of fluorescence intensity at the donor site is monitored in a macroscopic assembly of molecules. A distribution of decay rates is obtained, and as the decay rate depends on the distance between donor and acceptor, this can be used to infer the probability distribution of distances separating the donor and acceptor sites in the ensemble (*SI Appendix*).

In this study, a multisite, TR-FRET methodology coupled to MEM analysis has been used to study the time evolution of the probability distributions of 4 intramolecular distances in the small protein monellin, as it unfolds in 4 M guanidine hydrochloride (GdnHCl), starting from the native state. Single chain monellin (MNEI) is a sweet plant protein, whose folding and unfolding reactions have been studied extensively (26–28). Here, 4 single cysteine, single tryptophan containing mutant forms of MNEI have been used (Fig. 1*A*), in each of which a single tryptophan residue at position 4 in the sequence, Trp4, served as the donor fluorophore (D), and a thionitrobenzoate (TNB) adduct attached to a differently located thiol (at the positions shown in Fig. 1*A*) served as the FRET acceptor (A). It is shown that for all 4 structurally very similar proteins, the unfolding reaction can be characterized as slow diffusive swelling, indicating that the protein spends substantial time in many intermediate conformations, before unfolding completely.

Results and Discussion

Quenching of Fluorescence of Trp4 in Presence of TNB Is Due to FRET. The secondary structures and the stabilities of the unlabeled and TNB-labeled mutant variants are similar to that of wild-type (WT) MNEI (Cys42) (Fig. S1 and Table S1). All of the mutant variants exhibit identical fluorescence emission spectra, in the N and in the U states of the unlabeled protein. The quenching of the fluorescence of Trp4 in TNB-labeled Cys42 by FRET is apparent in (*i*) a reduction in fluorescence intensity at all wavelengths (Fig. 1*B*), with a larger reduction in N than in U, because Trp4 and C42-TNB are more separated in the latter. (*ii*) A shift in the position of the peak of the fluorescence lifetime distribution in the N state from 2.6 to 0.06 ns (Fig. 1*C*). For the other TNB-labeled proteins too, the fluorescence lifetime of Trp4 is reduced by different amounts, in both the N and the U states (Table S2), indicating that the quenching is distance-dependent and, hence, due to FRET. As expected for distance-dependent quenching, the mean fluorescence lifetime (τ_m) of Trp4 increases upon unfolding, for all of the TNB-labeled

Author contributions: S.K.J. and J.B.U. designed research; S.K.J. performed research; G.K. and J.B.U. contributed new reagents/analytic tools; S.K.J., D.D., and G.K. analyzed data; and S.K.J., D.D., G.K., and J.B.U. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. E-mail: ddhar@theory.tifr.res.in, gk@tifr.res.in, or jayant@ncbs.res.in.

This article contains supporting information online at www.pnas.org/cgi/content/full/0812564106/DCSupplemental.

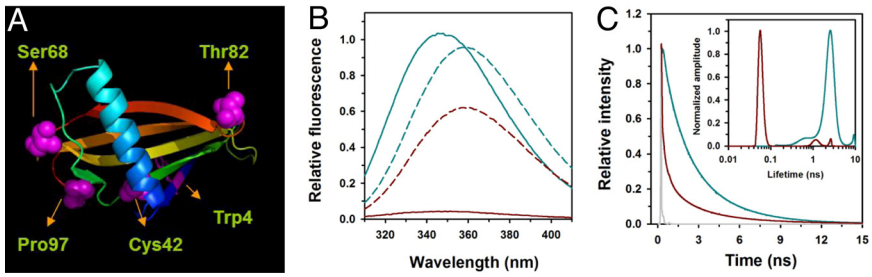


Fig. 1. FRET as a measure of intramolecular distances. (*A*) The location of Trp4 and the residues that were replaced by cysteine residues in the single Cys, single Trp-containing mutant variants are shown in the structure of single chain monellin (PDB entry 1IV7), which has been drawn using the program PyMOL (29). Trp4 is partially solvent accessible in the native state and the thiol label, TNB, quenches the fluorescence of Trp4 in a distance-dependent manner. (*B*) Fluorescence spectra of unlabeled native (solid blue line); TNB-labeled native (solid dark red line); unlabeled unfolded (dashed blue line); and TNB-labeled unfolded (dashed dark red line) Cys42. Trp4 shows maximum fluorescence emission at 346 nm in N, which shifts to 357 nm upon unfolding in 4 M GdnHCl. (*C*) Fluorescence intensity decay kinetics of unlabeled native (solid blue line) and TNB-labeled native (solid dark red line) Cys42. (*Inset*) The fluorescence lifetime distributions of unlabeled native (solid blue line) and TNB-labeled native Cys42 (solid dark red line) obtained from MEM analysis. Note the logarithmic *x* axis.

proteins, whereas it is not altered much (<5%) for the unlabeled proteins (Table S2). More specifically, the principal (≥84%) component of the fluorescence lifetime contributing to the τ_m also remains unchanged for all of the unlabeled proteins (Table S2). The mean value of the Forster's distance, R_0 (*SI Appendix*), for the Trp-TNB FRET pair, in the different variants of MNEI, was determined to be 22.8 Å (Table S3), making it an ideal probe to measure distances in the range of ≈10 to 38 Å, in the TNB-labeled proteins (30) with precision. The fluorescence lifetime measurements recover the expected D–A distances in the N states for all of the TNB labeled proteins (*SI Appendix* and Tables S2 and S3), confirming that the quenching is due to FRET.

Results from Ensemble-Averaging Probes Are Consistent with Apparent Two-State Unfolding. For each mutant protein, the fluorescence lifetime distributions and intensities in the TNB-labeled form were determined at different unfolding times, using a stopped-flow mixing module attached either to a fast time-correlated single photon counting (TCSPC) system (Fig. S2), or to a fluorimeter. For each protein, the time courses of the fractional increase in τ_m and fluorescence intensity were identical (Fig. 2 and Table S4), as expected, validating the accuracy of the TCSPC data collection in the double-kinetic experiments. The circular dichroism (CD) monitored kinetics were also identical to those monitored by FRET (Fig. S3). For all proteins, and for all probes, the time dependence of the entire expected change in signal is well-fitted by a simple exponential dependence (Fig. 2 *Insets*). Hence, the results of the integrated fluorescence intensity and CD measurements are consistent with a simple apparently two-state unfolding reaction, but a single time-dependent ensemble-averaged value of the observables does not reveal anything about the heterogeneity within the ensembles of structures present at any time during the reaction.

D–A Distances Change Continuously During the Time-Course of Unfolding. The complexity of the unfolding reaction is revealed when the fluorescence decay kinetics curves determined at various times of unfolding of the TNB-labeled proteins, were analyzed using MEM (Fig. 3). The fluorescence lifetime distributions so obtained were transformed into probability distributions of D–A distances (Fig. 4) using Forster's equation (*SI Appendix*, Eq. **S3**). Before unfolding commences, the N state of each protein displays a narrow fluorescence lifetime distribution, and hence, a narrow distribution of D–A distances (Figs. 3 and 4), which is different for the different mutant proteins. A narrow distribution is what is expected for a homogeneous N state. Once unfolding commences, the probability distribution of fluorescence lifetimes is bimodal at intermediate times for 3 mutant proteins, but appears unimodal at all times for the fourth mutant protein. The bimodal fluorescence lifetime distributions seen for Cys42-TNB, Cys82-TNB and Cys97-TNB give rise to bimodal distance distributions, and the unimodal fluorescence lifetime distribution seen for Cys68-TNB gives rise to a unimodal distance distributions, at intermediate times of unfolding (Figs. 3 and 4).

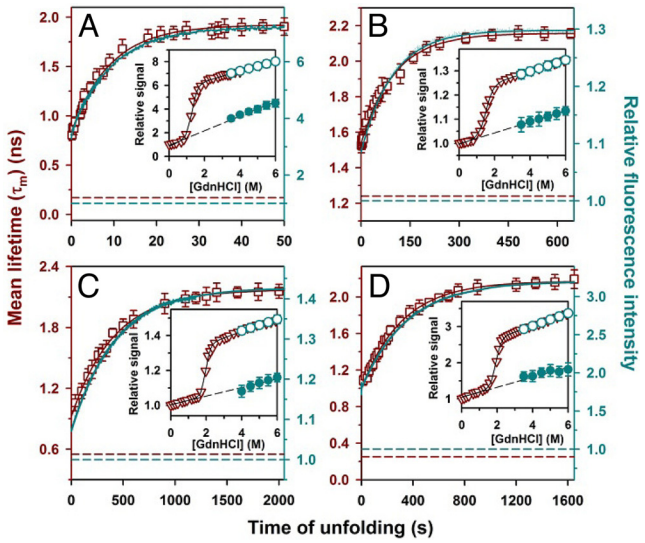


Fig. 2. Kinetics of unfolding of TNB-labeled mutant variants of MNEI in 4 M GdnHCl at pH 8 and 25 °C as monitored by steady-state and time-resolved FRET for Cys42-TNB (*A*), Cys68-TNB (*B*), Cys82-TNB (*C*), and Cys97-TNB (*D*). The change in the mean fluorescence lifetimes, τ_m ($\tau_m = \Sigma\alpha_{fi}\tau_{fi}$; where α_{fi} is the relative amplitude corresponding to relaxation time τ_{fi}) of Trp4 during unfolding (dark red open squares) are shown according to the left *y* axis. The solid dark red lines are a fit of the τ_m data to a single exponential equation; the dashed dark red lines represents the τ_m of the native protein in 0 M GdnHCl. Error bars represent the standard deviations of measurements from 3 different experiments. Blue solid lines show the changes in the fluorescence intensity of Trp4 during unfolding according to the right *y* axis. Dashed blue lines represent the fluorescence intensity of the native protein in 0 M GdnHCl. (*Insets*) The kinetic amplitudes of unfolding match the equilibrium amplitudes. The dark red inverted triangles represent the equilibrium unfolding transition monitored by the change in fluorescence intensity at 360 nm, and the continuous lines through the data represent a fit to a two-state N ⇄ U model. The blue open circles represent the $t = \infty$ signal, and the blue filled circles represent the $t = 0$ signal respectively, obtained from fitting the kinetic traces of unfolding to a single exponential equation. The black dashed lines are a linear extrapolation of the native protein baseline.

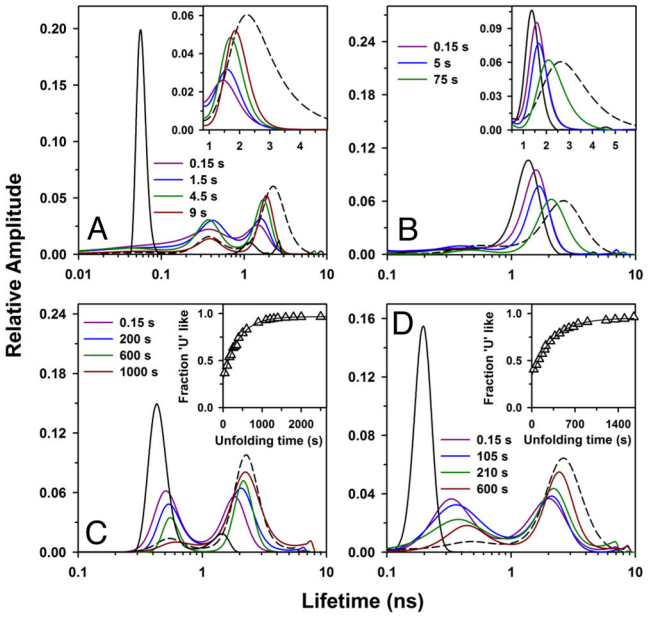


Fig. 3. Fluorescence lifetime distributions of the TNB-labeled mutant variants of MNEI at different times of unfolding. (*A*) Cys42-TNB. (*B*) Cys68-TNB. (*C*) Cys82-TNB. (*D*) Cys97-TNB. Black solid and dashed lines represent the fluorescence lifetime distributions of the native protein in 0 M GdnHCl and the unfolded protein in 4 M GdnHCl, respectively. (*A* and *B Insets*) Gradual evolution of the fluorescence lifetime distributions, with the time of unfolding, on a linear scale. (*C* and *D Insets*) Fractions of protein present (open triangles) in U-like forms (those whose lifetimes are similar to those of the U distribution) at different unfolding times, as determined from the TR-FRET measurement of fluorescence lifetime distributions. The fraction was calculated by dividing the area under the U-like peak of the bimodal lifetime distributions, at various unfolding times, by the total area under the U distribution obtained at the end of the unfolding reaction in 4 M GdnHCl. It was possible to do this for Cys82-TNB and Cys97-TNB because in these cases the position of the fluorescence lifetime distribution of the unfolded protein changes only marginally with the time of unfolding.

The unfolding reactions of Cys42-TNB, Cys82-TNB and Cys97-TNB show the following features (Figs. 3 and 4): (*i*) At the first time of measurement, 150 ms, when less than a 1% change in CD has occurred, the largely unimodal fluorescence lifetime and distance distributions of N, have transformed into broad bimodal distributions. Clearly, the ensemble of conformations formed at this time, although still structured, is different from the N state. The 2 peaks are centered at distances intermediate between, and markedly different from, those of N and U (Tables S2 and S3), which implies the existence of 2 different intermediate structures, I_1 and I_2, populated within 150 ms of the commencement of unfolding. (*ii*) I_1 and I_2 display much broader distributions than does N, indicating that these intermediates are not unique structures but are ensembles of structures. I_1 shows more native-like distance distributions, although in the case of the Trp4-C42-TNB distance, there is practically no overlap with the N distribution. The peaks corresponding to I_2 are closer to those of U than I_1, but distinct from the U distributions in all 3 proteins, more so in the case of Cys42-TNB. (*iii*) As unfolding proceeds, the peak corresponding to I_1 shifts either not at all (Cys42-TNB and Cys82-TNB) or marginally (Cys97-TNB), to larger distances. (*iv*) The peak corresponding to I_2 evolves gradually into those of U, indicating a continuum of intermediate structures from I_2 to U. (*v*) The population of I_2 increases continuously at the expense of that of I_1.

In the case of both Cys82-TNB and Cys97-TNB, the kinetics of the change in the area under the peak corresponding to I_2 (see legend to Fig. 3), are similar to that of the change in the fraction of U calculated from the changes in fluorescence intensity (Fig. 2) or CD (Fig. S3), assuming that only N and U are present at any time during unfolding. This result validates the use of TR-FRET experiments coupled to MEM analysis, to estimate quantitatively the populations of different conformational states during unfolding.

The changes in the fluorescence lifetime and distance distributions during the unfolding of Cys68-TNB appear to be different from those seen during the unfolding of the other mutant proteins. At 150 ms of unfolding, a unimodal distribution (see legend to Fig. 4) similar to that of N, but shifted to a slightly larger distance, is seen. As unfolding proceeds, the fluorescence lifetime and distance distributions shift gradually to the distribution of U (Figs. 3*B* and 4), remaining unimodal at all times. It should be noted that at no time of unfolding, can the distributions be fitted to a linear combination of the distributions obtained for N and U states. Nor can the gradual increase in fluorescence lifetime and distance be explained by very rapid exchange between N and U, because (*i*) the rate of exchange would have to be faster than 10^{10} s^{-1}, the fastest rate of fluorescence decay, whereas conformational changes in proteins (31) are found typically to be slower than 10^9 s^{-1}, and (*ii*) the observed changes in the widths of the distributions (Figs. 3 and 4) with the time of unfolding cannot be explained by the fast exchange model. Hence, it is concluded that the Trp4-C68-TNB distance expands gradually during unfolding.

Protein Swelling During Unfolding. The experimental data on the time evolution of the probability distribution of intramolecular distances during unfolding, can be understood in terms of a simple phenomenological model, which is a variation of the well-known Rouse model of polymer physics (32, 33). The model (Scheme 1; see *SI Appendix* for details) is based on the following observations and assumptions: (*i*) Within 150 ms of unfolding, the conformation of any protein molecule is observed to be close to 1 of 2 intermediate ensembles of structures, called I_1 and I_2 (Figs. 3 and 4). A few non-nearest-neighbor, noncovalent bonding interactions appear to persist in the more compact I_1, and fewer still in the more swollen I_2. It is likely that I_1 and I_2 represent the products of 2 parallel unfolding reactions, in which many intrachain noncovalent interactions break in different parts of the chain. In fact, in an earlier study of the unfolding of MNEI (27), native protein had been observed to transform into differentially unfolded forms (15% and 85%) within a few ms of unfolding. Here, it appears that I_1 and I_2 form with probability of 0.2 and 0.8, respectively (Figs. 3 and 4). (*ii*) During the slow swelling with time within the ensembles of states I_1 and I_2, the probability distribution function of D–A distances is well approximated by a Gaussian probability distribution function,

$$P(R,t) = \frac{1}{\sqrt{2\pi}\sigma_i(t)} e^{\frac{(R-R_i(t))^2}{2\sigma_i^2(t)}}$$

$R_i(t)$ and $\sigma_i^2(t)$ are the time-dependent mean and variance, respectively, of the distance distribution in the ensemble I_i. Although the Gaussian form is known to be incorrect for very small or very large R value, this simple form has been used because it provides adequate qualitative fits to the data (Fig. 4), and more accurate fitting forms would require more parameters. It should be noted that although I_1 and I_2 possess structure (see below), simulations have shown that the end-to-end distance of a structured polypeptide chain is approximately Gaussian-distributed (34). In N itself, the 4 distances appear to be Gaussian-distributed (Fig. 4). (*iii*) The molecules in I_1 and I_2 swell with time of unfolding (Fig. 4), with characteristic times τ_1 and τ_2, respectively. Swelling is assumed to occur in a succession of small local and random conformational changes. (*iv*) The

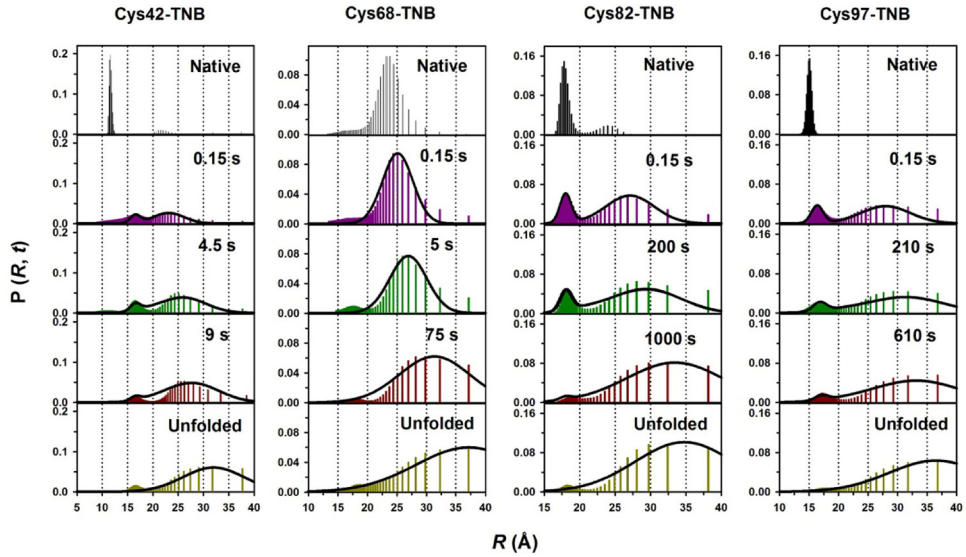
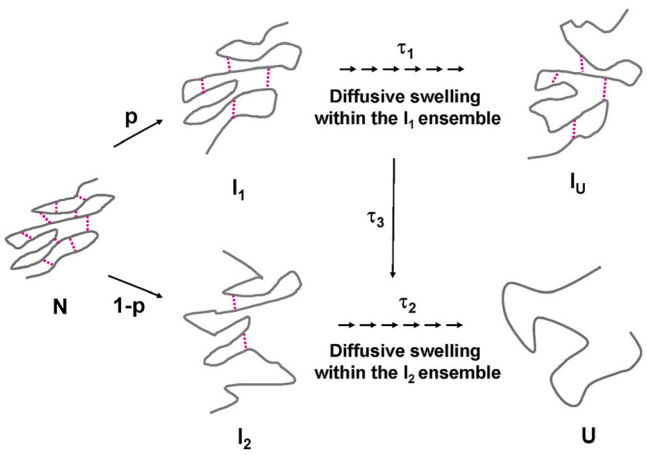


Fig. 4. Evolution of the probability distributions of D–A distances with time of unfolding. The fluorescence lifetime distributions at different times of unfolding were converted into distance distributions as described in *SI Appendix*. The 22.8 Å Forster's distance for the Trp-TNB pair does not allow accurate measurement of the D–A distance beyond ≈38 Å. Consequently, the single bars at ≈38 Å in the distributions represent the population of polypeptides with $R \geq 38$ Å. In the case of Cys68-TNB, 3 - 4% of the population appears to be distributed around a distance less than that in N, at different times of unfolding. This appears to be an experimental artifact of unknown origin. It cannot represent the population distribution of I_1, which in the case of the other proteins, is centered at a distance intermediate between that of N and U. The distance distributions were also calculated using Eq. **S16**, which was derived on the basis of the simple phenomenological model described in *SI Appendix*. The black continuous curves shows the calculated distance distributions at the indicated times of unfolding. The values of the parameters used for generating the distance distributions are listed in Table S5.

qualitative features of the swelling of a real protein chain can be described adequately by a Rouse-like model. The physics of swelling of the polypeptide chain in both I_1 and I_2, upon unfolding, is assumed to be similar to that of a freely jointed self-avoiding chain, with a few additional non-nearest-neighbor, noncovalent interactions. For Rouse-like relaxational dynamics (32), the polymer, starting from a N-like conformation, would slowly swell with time, and its mean size would increase with time to a finite limiting value. The microscopic relaxation time (Rouse time) (32) corresponding to the rotational relaxation time of the protein chain, which is ≈4 ns for monellin (*SI Appendix*), is much shorter than the time scale of the unfolding reaction. In the Rouse-like model, the second longest relaxation time is approximately a fourth of the longest relaxation time. Hence, for large times, the time dependence of $R_i(t)$ and $\sigma_i^2(t)$ can be approximated by a simple exponential, with relaxation time τ_i. The fractional increases in $\sigma_i^2(t)$ and $R_i(t)$ are therefore the same at any time t (*SI Appendix*). In the case of homopolymer swelling too, a simple physical model based on using the Langevin equation to describe polymer dynamics, has shown that the polymer size tends to its finite limiting value exponentially with time, for times much greater than the Rouse time (35). (*v*) The population of I_2 is observed to grow slowly at the expense of I_1 (Fig. 4), and this transformation is assumed to be a first order process with time constant τ_3. (*vi*) I_2 swells slowly to finally form U.



Scheme 1. Diffusive swelling during unfolding.

Fig. 4 shows that this simple model can fit the data reasonably well. It is remarkable that the very different behaviors of the 4 distance distributions measured during unfolding, can be attributed fully just to differences in the values of the various kinetic parameters defining Eq. **S16** (values used for fitting are listed in Table S5). In particular, it is seen that the distance distributions measured at different times of unfolding of Cys68-TNB, appear to be unimodal, unlike those of the other proteins, only because the distributions in I_1 and I_2 cannot be resolved. Importantly, the values of τ_1 and τ_2 are found to be different for the different TNB-labeled mutant proteins (Tables S5 and S6). This indicates that the energies of different intramolecular interactions in the molecule are not unaffected by the position of the FRET acceptor. Here, it should be mentioned that the expansion of an intramolecular distance is observed to be slower for a distance separating more amino acid residues in the polypeptide chain, but no scaling exponents can be estimated reliably from the data. It should be noted that even though Cys42-TNB is less stable, and hence unfolds faster than the other labeled proteins (Figs. S1 and S3), this in no way vitiates the basic result that in all 4 structurally very similar (Fig. S1) mutant proteins, clear evidence of slow diffusive swelling is observed.

The Rouse-like model describes well, in a qualitative manner, the observed diffusive swelling of the protein chain. Effects like excluded volume, are not included in the simple model used here. Although these effects would change some exponents used in Eq. **S16**, the deviations of these exponents from simple Rouse values are not detectable in the data (Fig. 4). It should be noted that the Rouse model is only a zeroth level approximation to the true dynamics of the protein chain, and cannot describe fully the

details of the swelling process. For example, the attractions within the protein chain and with the solvent will be important determinants of the true dynamics.

One feature of the data that cannot be explained by the simple phenomenological model used here is why the unfolding times are so long. For a Rouse-like chain of N monomers, the relaxation time varies as $\tau_0 N^z$. With $\tau_0 \sim 1$ ps (for the basic flip process in Rouse-like dynamics), and $z \sim 2$ to 3, a relaxation time substantially smaller than the observed time scale, is predicted. A possible explanation is that the activation barrier for the basic flip process is substantially larger in the more compact I_1 and I_2 than in the more expanded U, because of additional steric hindrance due to the presence of other parts of the chain nearby. In addition, packing (van der Waal's) and other (H-bonding) intramolecular interactions stabilizing the structures still present in I_1 and I_2, which distinguish them from U, would need to break during diffusive swelling. Hence, these interactions could play dominant roles in determining the chain swelling dynamics. Another possible reason could be that unfolding is slowed down by many small ($\approx$3–4 k_BT) free energy barriers distributed over a multidimensional free energy surface: Diffusive swelling is likely to be retarded by hindered dynamics, and even backtracking (36) through a multitude of local energy minima (3–6, 37). Ruggedness in the energy landscape would arise because of imperfect enthalpy-entropy compensation as intraprotein and protein-solvent interactions break throughout the unfolding process, and because of entropic barriers associated with the immobilization of water molecules on hydrophobic surfaces that become exposed. There is some evidence from computer simulations that the diffusive motions of a polypeptide chain at large times over a rugged energy landscape, can still give rise to relaxation that is nearly exponential (38), and the data presented here provide the first direct experimental evidence that a gradual unfolding process can also show nearly exponential relaxation in some ensemble-averaged observables.

Specific Structure Is Lost Gradually During the Unfolding of MNEI. Earlier studies of the folding and unfolding of monellin (26–28) had shown that both processes are complex and mediated by multiple compact intermediates. This study has further brought out the complexity of the unfolding reaction by showing that the initially formed intermediate ensembles, I_1 and I_2, unfold gradually, and not just in a few steps as indicated in the study in ref. 27. Here, it is shown that I_1 and I_2 possess native-like structure, as judged by the observation of only a very marginal loss in secondary structure at 150 ms of unfolding (Fig. S3). Thus, the important result of this study is not only that I_1 and I_2 can be observed to swell and unfold in a continuous manner, but also that nearly all specific structure is lost during the course of this slow gradual unfolding process. I_1 and I_2 are seen to be expanded forms of the protein, in which some of the non-nearest-neighbor interactions have been lost. In this respect, they resemble the dry molten globule state, which has been predicted by theory (10), and which has also been observed to form very early during unfolding, both by experiment (39–41) and simulation (42). The main conclusion of this study, that unfolding may occur via a continuum of progressively less structured conformations, can also have implications for understanding protein aggregation reactions, such as amyloid fibril formation, because such aggregation can commence from different conformations of partially unfolded proteins (43).

Methods

Protein Expression, Purification, and TNB-Labeling. WT MNEI contains a single tryptophan (Trp4) and a single cysteine (Cys42) residue. In addition to WT MNEI (Cys42), 3 more mutant variants, C42AS68C (Cys68), C42AT82C (Cys82) and C42AP97C (Cys97), each with a single tryptophan residue (Trp4) and a single cysteine residue were generated by site-directed mutagenesis, and the proteins purified as described in ref. 27. The TNB-labeled proteins were obtained as described in *SI Appendix*.

Steady-State Fluorescence and Far-UV CD-Monitored Equilibrium and Kinetic Unfolding Experiments. The secondary structure, stabilities and kinetics of unfolding of all of the unlabeled and TNB-labeled proteins were measured as described in *SI Appendix*.

TR-FRET Monitored Unfolding Kinetics Experiments. Fluorescence lifetimes were measured at different times of the unfolding reaction for all of the unlabeled and TNB-labeled proteins. For this purpose a stopped-flow module (RX2000 rapid kinetics spectrometer accessory from Applied Photophysics) was synchronized with a femto/pico second Ti-sapphire laser (Spectra Physics) and a dual channel fast time-correlated single photon counting (TCSPC) system (Fig. S2). One-picosecond pulses of 885-nm radiation from the Ti-sapphire laser, pumped by an Nd:YUO, laser (Millenia X, Spectra Physics), were frequency tripled to 295 nm by using a frequency doubler/tripler (GWU, Spectra physics). The resultant 295-nm laser radiation was used as the excitation source and focused on to the sample in the cuvette coupled to the stopped flow module. The unfolding reaction was initiated by mixing the native protein with the unfolding buffer inside the cuvette of the stopped-flow mixer in the ratio, 1:5.5. The dead-time of the mixing was estimated to be $\approx$60 ms. At different times of the unfolding process, the fluorescence intensity decay transients were collected at 360 nm, after passing the emitted light through a 320-nm high pass cut-off filter, to avoid the collection of any scattered excitation light. The fluorescence decay data collection was done using data acquisition windows of $\approx$100-150 ms, using either 1 or 2 microchannel plate photomultiplier tubes (Model H5773P-01; Hamamatsu) coupled to a TCSPC card (SPC-630; Becker & Hickl), which was operated in the reverse start-stop mode. The instrument response function (IRF) at 295 nm was obtained using a dilute colloidal suspension of dried non-dairy coffee whitener. The emission was monitored at the magic angle (54.7°) to eliminate the contribution from the decay of anisotropy. Care was taken to ensure that the peak counts in the fluorescence decays were at least 20,000.

Analysis of Time-Resolved Fluorescence Intensity Decays. The fluorescence decay curves were subjected to discrete analysis and MEM analysis, as described in refs. 44 and 45. See *SI Appendix* for details of discrete analysis. In MEM analysis, it is assumed that the decay originates from a distribution of fluorescence lifetimes in the range 10 ps to 10 ns, or in a similar range, with all lifetimes initially assumed to have equal probability weights (amplitude). Subsequently, the distribution is modified in each iteration of the analysis, leading to a maximization of the Shannon–Jaynes entropy, $S = \Sigma - P_i \log P_i$, where P_i is the probability (amplitude) of the ith lifetime, and to a minimization of the residuals ($\chi 2$). For a particular value of χ^2, there could be many possible values of P_i. MEM analysis identifies the distribution for which S is maximum. The calculation is terminated when successive iterations do not change the values of χ^2, S, and the distribution profile. Thus, MEM analysis gives the best (least biased) fit to the data without assuming any particular line-shape function.

Confirmation of MEM Analysis. The robustness of the lifetime distributions (including peak positions and widths of distributions) obtained by MEM analysis was checked exhaustively by collecting data on several samples under the same sample conditions. Peak values of MEM distributions agreed with those obtained from discrete lifetime analyses, within $\approx$5% (Fig. S4). Values of χ^2 were in the range of 0.95 to 1.15 for all of the discrete and MEM analyses. See *SI Appendix* for additional details.

Analysis of FRET and Calculations of Distances. Fluorescence lifetimes were converted into D–A distances using Forster's equation, as detailed in *SI Appendix*.

Generation of Distance Distributions. Distance distributions were also generated using the phenomenological model described in detail in *SI Appendix*, using the software package Mathematica, Version 6.

ACKNOWLEDGMENTS. We thank A. K. Patra for generating the mutant variants of MNEI used in this study; N. Periasamy for the MEM software and advice on its use, and members of our laboratory for discussions and comments on the manuscript. This work was supported by a Shyama Prasad Mukherjee Fellowship awarded by the Council of Scientific and Industrial Research, India (to S.K.J.); J. C. Bose National Research Fellowships from the Government of India (to D.D. and J.B.U.); the Tata Institute of Fundamental Research; and the Department of Science and Technology, Government of India.

1. Scheraga HA (1971) Theoretical and experimental studies of conformations of polypeptides. *Chem Rev* 71:195–217.
2. Dill KA (1985) Theory for the folding and stability of globular proteins. *Biochemistry* 24:1501–1509.
3. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins* 21:167–195.
4. Dill KA, Chan HS (1997) From Levinthal to pathways to funnels. *Nat Struct Biol* 4:10–19.
5. Chavez LL, Onuchic JN, Clementi C (2004) Quantifying the roughness on the free energy landscape: Entropic bottlenecks and protein folding rates. *J Am Chem Soc* 126:8426–8432.
6. Frauenfelder H, Fenimore PW, Young RD (2007) Protein dynamics and function: Insights from the energy landscape and solvent slaving. *IUBMB Life* 59:506–512.
7. Klimov DK, Thirumalai D (2000) Mechanisms and kinetics of β-hairpin formation. *Proc Natl Acad Sci USA* 97:2544–2549.
8. Klimov DK, Thirumalai D (2002) Is there a unique melting temperature for two-state proteins? *J Comput Chem* 23:161–165.
9. Karplus M (2000) Aspects of protein reaction dynamics: Deviations from simple behavior. *J Phys Chem B* 104:11–27.
10. Finkelstein AV, Shakhnovich EI (1989) Theory of cooperative transitions in protein molecules. II Phase diagram for a protein molecule in solution. *Biopolymers* 28:1681–1694.
11. Englander SW (2000) Protein folding intermediates and pathways studied by hydrogen exchange. *Annu Rev Biophys Biomol Struct* 29:213–238.
12. Swaminathan R, Nath U, Udgaonkar JB, 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.
13. Holtzer ME, Lovett EG, d'Avignon DA, Holtzer A (1997) Thermal unfolding in a GCN4-like leucine zipper: $^{13}C^{\alpha}$ NMR chemical shifts and local unfolding curves. *Biophys J* 73:1031–1041.
14. Lakshmikanth GS, Sridevi K, Krishnamoorthy G, Udgaonkar JB (2001) Structure is lost incrementally during the unfolding of barstar. *Nat Struct Biol* 8:799–804.
15. Kuzmenkina EV, Heyes CD, Nienhaus GU (2006) Single-molecule FRET study of denaturant induced unfolding of RNase H. *J Mol Biol* 357:313–324.
16. Sadqi M, Fushman D, Munoz V (2006) Atom-by-atom analysis of global downhill protein folding. *Nature* 442:317–321.
17. Li H, Frieden C (2007) Comparison of C40/82A and P27A C40/82A barstar mutants using ^{19}F NMR. *Biochemistry* 46:4337–4347.
18. Gruebele M (2005) Downhill protein folding: Evolution meets physics. *C R Biol* 328:701–712.
19. Sinha KK, Udgaonkar JB (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.
20. Sinha KK, Udgaonkar JB (2008) Barrierless evolution of structure during the submillisecond refolding reaction of a small protein. *Proc Natl Acad Sci USA* 105:7998–8003.
21. Beechem JM (1997) Picosecond fluorescence decay curves collected on millisecond time scale: Direct measurement of hydrodynamic radii, local/global mobility, and intramolecular distances during protein-folding reactions. *Methods Enzymol* 278:24–49.
22. Ratner V, Sinev M, Haas E (2000) Determination of intramolecular distance distribution during protein folding on the millisecond timescale. *J Mol Biol* 299:1363–1371.
23. Lyubovitsky JG, Gray HB, Winkler JR (2002) Mapping the cytochrome C folding landscape. *J Am Chem Soc* 124:5481–5485.
24. Sridevi K, Lakshmikanth GS, Krishnamoorthy G, Udgaonkar JB (2004) Increasing stability reduces conformational heterogeneity in a protein folding intermediate ensemble. *J Mol Biol* 337:699–711.
25. Udgaonkar JB (2008) Multiple routes and structural heterogeneity in protein folding. *Annu Rev Biophy* 37:489–510.
26. Kimura T, et al. (2005) Specific collapse followed by slow hydrogen-bond formation of β-sheet in the folding of single-chain monellin. *Proc Natl Acad Sci USA* 102:2748–2753.
27. Patra AK, Udgaonkar JB (2007) Characterization of the folding and unfolding reactions of single-chain monellin: Evidence for multiple intermediates and competing pathways. *Biochemistry* 46:11727–11743.
28. Kimura T, et al. (2008) Dehydration of main-chain amides in the final folding step of single-chain monellin revealed by time-resolved infrared spectroscopy. *Proc Natl Acad Sci USA* 105:13391–13396.
29. DeLano WL (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, *USA*.
30. Saxena AM, Udgaonkar JB, Krishnamoorthy G (2006) Characterization of intramolecular distances and site-specific dynamics in chemically unfolded barstar: Evidence for denaturant-dependent non-random structure. *J Mol Biol* 359:174–189.
31. Ishima R, Torchia DA (2000) Protein dynamics from NMR. *Nat Struct Biol* 7:740–743.
32. Doi M (2001) In *Introduction to Polymer Physics* (Oxford Univ Press, Oxford), pp 65–88.
33. Dua A, Adhikari R (2008) Non-Markovian fluctuations in Markovian models of protein dynamics. arXiv:0810.2232v1.
34. Fitzkee NC, Rose GD (2004) Reassessing random-coil statistics in unfolded proteins. *Proc Natl Acad Sci USA* 101:12497–12502.
35. Pitard E, Orland H (1998) Dynamics of the swelling or collapse of a homopolymer. *Europhys Lett* 41:467–472.
36. Gosavi S, Chavez LL, Jennings PA, Onuchic JN (2006) Topological frustration and the folding of interleukin-1 β. *J Mol Biol* 357:986–996.
37. Zwanzig R (1988) Diffusion in a rough potential. *Proc Natl Acad Sci USA* 85:2029–2030.
38. Hagen SJ (2003) Exponential decay kinetics in ''downhill'' protein folding. *Proteins* 50:1–4.
39. Kiefhaber T, Labhardt AM, Baldwin RL (1995) Direct NMR evidence for an intermediate preceding the rate-limiting step in the unfolding of ribonuclease A. *Nature* 375:513–515.
40. Hoeltzli SD, Frieden C (1995) Stopped-flow NMR spectroscopy: Real-time unfolding studies of 6-^{19}F-tryptophan-labeled *Escherichia coli* dihydrofolate reductase. *Proc Natl Acad Sci USA* 92:9318–9322.
41. Rami BR, Udgaonkar JB (2002) Mechanism of formation of a productive molten globule form of barstar. *Biochemistry* 41:1710–1716.
42. Hua L, Zhou R, Thirumalai D, Berne BJ (2008) Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. *Proc Natl Acad Sci USA* 105:16928–16933.
43. Calamai M, Chiti F, Dobson CM (2005) Amyloid fibril formation can proceed from different conformations of a partially unfolded protein. *Biophys J* 89:4201–4210.
44. 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.
45. Swaminathan R, Periasamy N (1996) Analysis of fluorescence decay by 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.