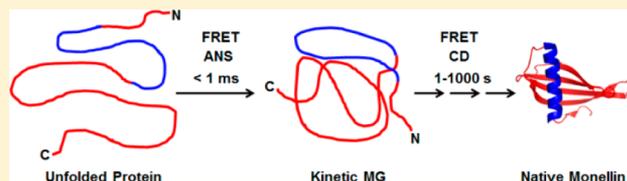


# Rise of the Helix from a Collapsed Globule during the Folding of Monellin

Rama Reddy Goluguri and Jayant B. Udgaonkar\*

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

**ABSTRACT:** Early kinetic intermediates observed during the folding of many proteins are invariably compact and appear to possess some secondary structure. Consequently, it has been difficult to understand whether compaction drives secondary structure formation or secondary structure formation facilitates compaction during folding. In this study of the folding of single-chain monellin, it is shown that a kinetic molten globule (MG) is populated at 2 ms of folding. Far-UV circular dichroism (CD) measurements show that the kinetic MG is devoid of any helical structure even under the most stabilizing folding conditions. Multisite fluorescence resonance energy transfer (FRET) measurements show that the kinetic MG is compact with different segments having contracted to different extents. It is shown that the sequence segment that goes on to form the sole helix in the native protein is fully collapsed in the kinetic MG. This segment expands to accommodate the helix as the kinetic MG folds further to the native state, while other segments of the protein contract. Helix formation starting from the kinetic MG is shown to occur in multiple kinetic steps, whether measured by far-UV CD or by FRET.



Little is understood about how the ubiquitous molten globule (MG), found to form early on the folding pathways of many proteins, folds into the structured native (N) state of the protein. MGs formed by a wide variety of proteins have been shown to be compact<sup>1–3</sup> but without a well-packed core;<sup>4,5</sup> to possess some degree of secondary structure,<sup>6</sup> although this structure may not be stable<sup>7,8</sup> and may not even be natively-like;<sup>9–13</sup> and to possess very few specific tertiary interactions but yet have a natively-like topology.<sup>14,15</sup> The MG appears to be a heterogeneous ensemble of different conformations whose stability and composition, and hence structure, depend upon the stabilization conferred by the specific conditions under which folding occurs.<sup>16–19</sup> MGs populated on folding pathways (kinetic MGs) are important to study because the major structural features of a native protein, tight packing and significant specific structure, appear to arise only after the kinetic MGs first form.<sup>20–22</sup> It is likely that the partial structure that a kinetic MG does possess acts as the seed for the major structure formation reactions that progressively transform it to the N state.

Kinetic MGs typically form in <1 ms. During their formation, it appears that specific secondary structure formation is preceded by a nonspecific collapse of the polypeptide chain.<sup>16,23–25</sup> This is not surprising given that polypeptide chain collapse reactions can occur in times as short as 50 ns,<sup>26,27</sup> and the formation of secondary structures, at least in isolated peptides, is typically somewhat slower.<sup>28</sup> Moreover, secondary structures in proteins invariably need to be stabilized by specific tertiary interactions.<sup>7,29,30</sup> Nevertheless, for some proteins, chain collapse and secondary structure formation have not been temporally resolved.<sup>31–33</sup> The difficulty in dissecting the kinetic process of MG formation into individual steps is compounded when the unfolded (U) state itself possesses

compact substructures,<sup>34,35</sup> including secondary structure,<sup>36,37</sup> which would seed further secondary structure formation in the kinetic MG. Indeed, in molecular dynamics simulations of the folding of several proteins, secondary structure was observed to form in the early stages of folding, presumably only because the starting U states already possessed secondary structures.<sup>38</sup> An understanding of how different segments of structure in the kinetic MG differ in their stabilities may not only provide a clue about how they form but also suggest how they are likely to develop further structure.

To understand how the compact nature of the kinetic MG can facilitate the formation of secondary structure, in particular helical structure, it is necessary to first populate a kinetic MG with little if any secondary structure. This becomes possible to do for some proteins under less stabilizing folding conditions,<sup>16</sup> but the kinetic MG invariably is found to possess secondary structure under folding conditions that confer greater stability.<sup>18,24,39</sup> MGs devoid of specific secondary structure are rare,<sup>40,41</sup> but once a kinetic MG devoid of specific structure (a structureless globule) is identified, it is also important to use a probe other than far-UV circular dichroism (CD) to probe for secondary structure formation, because aromatic residues,<sup>42,43</sup> as well as polyproline II conformations,<sup>44</sup> can also contribute significantly to the far-UV CD signal; hence, measurement of only far-UV CD may give erroneous results.

Fluorescence (Forster) resonance energy transfer (FRET) has been shown to be a powerful tool in the study of protein folding and unfolding reactions.<sup>16,23,41,45–55</sup> In FRET measurements, the change in distance separating a donor (D)

Received: June 29, 2015

Revised: August 4, 2015

Published: August 10, 2015

fluorophore from an acceptor (A) fluorophore during folding or unfolding is measured, and multisite FRET measurements allow structural changes in multiple segments of the protein structure to be probed. If an appropriate D–A fluorophore pair is placed suitably across a sequence stretch that will fold to form a helix, it should be possible to monitor the rise of the helix from a collapsed state, because the D–A distance would expand as the collapsed sequence stretch forms the rodlike helix.

The small protein, monellin, in its single-chain form, MNEI, is particularly well suited for studying the interplay of chain compaction and secondary structure formation during the early stages of folding. In its  $\beta$ -grasp fold, a single long helix is packed against a five-stranded  $\beta$ -sheet.<sup>56</sup> A kinetic MG has been shown to form within 1 ms of the initiation of its folding reaction.<sup>57,58</sup> The kinetic MG of MNEI is especially interesting to study because it has been shown to be heterogeneous with subpopulations from which different folding pathways operate in parallel.<sup>58,59</sup>

In this study, it is shown that the kinetic MG of MNEI is devoid of  $\alpha$ -helical secondary structure, as measured by far-UV CD, even under the most stabilizing folding conditions. A FRET D–A pair was placed at the ends of the sequence stretch spanning the helix that is present in the N state, and it is shown that this distance, as well as three other intramolecular distances also probed by FRET, has contracted substantially in the kinetic MG. The four distances have contracted to different extents, with the maximal contraction seen in the sequence stretch that will go on to form the helix. As the kinetic MG proceeds to fold, the distance spanning the latter sequence stretch expands to accommodate the helix, while the other segments of the protein become more compact.

## MATERIALS AND METHODS

**Buffers and Reagents.** All the chemicals used in this study were of high purity grade and obtained from Sigma. The guanidine hydrochloride (GdnHCl) used was of ultrapure grade and was obtained from USB chemicals. The native buffer used consisted of 50 mM phosphate (pH 7) containing 250  $\mu$ M EDTA and 1 mM DTT. For the labeled protein, DTT was not added to the buffer. The unfolding buffer used was native buffer containing 4 M GdnHCl. All buffers were filtered through 0.22  $\mu$ m filters and degassed before the experiment.

**Protein Expression and Purification.** The procedure for the purification of MNEI has been described previously.<sup>58</sup> The protein concentration was determined using an extinction coefficient of 14600 M<sup>-1</sup> cm<sup>-1</sup> at 280 nm.<sup>58</sup> In the case of the labeled protein, the absorbance at 280 nm was corrected for the contribution of the TNB label as described previously.<sup>50</sup>

**DTNB Labeling of MNEI.** The cysteine residue in different single-cysteine-containing mutant variants of MNEI was modified by adding the TNB moiety to it. The procedure for TNB labeling was the same as that described previously.<sup>53</sup> The extent of labeling was checked with electrospray ionization mass spectrometry, and the protein was found to be >95% labeled. Once labeling had occurred, the molecular weight increased by 196 Da, corresponding to the addition of a TNB moiety to the single cysteine residue.

**Equilibrium Unfolding Studies.** Equilibrium unfolding studies were conducted using a stopped flow module (SFM-4) from Biologic. The protein solution was mixed with buffer containing different concentrations of GdnHCl and incubated at room temperature for 6 h. The final protein concentration was 10  $\mu$ M. The samples were excited at 295 nm, and the

emission was monitored at 360 nm using a band-pass filter with a bandwidth of  $\pm 10$  nm (Asahi spectra).

CD-monitored equilibrium unfolding experiments were conducted using a JASCO J-815 spectropolarimeter. The ellipticity at 222 nm was measured using a cuvette with a path length of 2 mm and a bandwidth of 2 nm. The sample concentration used for CD measurement was 10  $\mu$ M. The signal was converted into relative ellipticity by normalizing all the signals to the signal of the native protein in 0 M GdnHCl.

**Fluorescence Emission Spectra.** The fluorescence spectra of the unlabeled and TNB-labeled proteins were acquired using a FluoroMax-3 (JobinYvon) spectrofluorimeter. The samples were excited at 295 nm, with an excitation bandwidth of 0.5 nm, and the emission was monitored from 310 to 450 nm with a bandwidth of 5 nm. Three scans were averaged for each sample. The protein concentration used was 10  $\mu$ M, and the path length of the cuvette was 1 cm.

**Stopped Flow Refolding Experiments.** All the millisecond refolding experiments were conducted on the SFM-4 instrument. A mixing dead time of 2 ms was achieved by using a flow rate of 5 mL/s with a cuvette with a path length of 0.8 mm. The protein was unfolded in 4 M GdnHCl for 4 h. Refolding was initiated by rapidly diluting the denaturant into a buffer containing different concentrations of GdnHCl.

For ANS-monitored refolding experiments, the denatured protein was diluted into refolding buffer containing ANS using the SFM-4 instrument and the MOS-450 acquisition system. The dead time of mixing was 6.2 ms. The sample was excited at 350 nm, and the emitted fluorescence was collected using a 450 nm band-pass filter with a bandwidth of 10 nm (Asahi spectra). The final ANS concentration in the reaction was 330  $\mu$ M, and the protein concentration was 10  $\mu$ M.

For CD-monitored refolding experiments, a SFM-400 module was interfaced with a JASCO J-815 spectropolarimeter. The dead time of mixing was 10 ms. The CD signal was monitored at 222 nm with a bandwidth of 2 nm. The final protein concentration used for stopped flow CD measurement was 25  $\mu$ M.

**Data Analysis. FRET Efficiency Calculation.** The FRET efficiencies for the different distances probed were calculated using the equation

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

where  $F_{DA}$  is the fluorescence of the donor in the presence of the acceptor, which is the value obtained from the labeled protein, and  $F_D$  is the fluorescence of donor alone, which is the value obtained from the unlabeled protein.

**Calculation of Fractional FRET Efficiencies.** The fractional change in FRET efficiency ( $F_E$ ) was calculated according to the equation

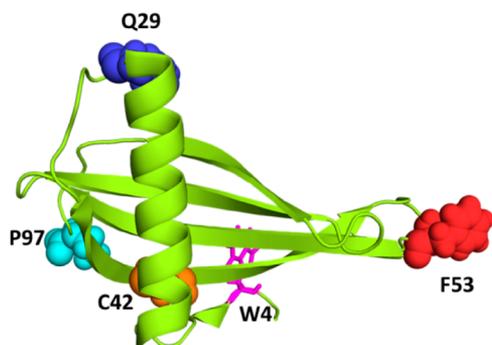
$$F_E = \frac{E_N - E_C}{E_N - E_U} \quad (2)$$

where  $E_N$  and  $E_U$  are the FRET efficiencies of the native and unfolded states, respectively, and  $E_C$  is the FRET efficiency of the collapsed species.

**Fitting the Equilibrium Unfolding Curves.** The equilibrium unfolding transition was fitted to a two-state N  $\leftrightarrow$  U model,<sup>60</sup> and the values of the free energy of unfolding in water ( $\Delta G_U$ ) and the slope of the transition ( $m$  value) were obtained from the fit using SigmaPlot12.

## RESULTS

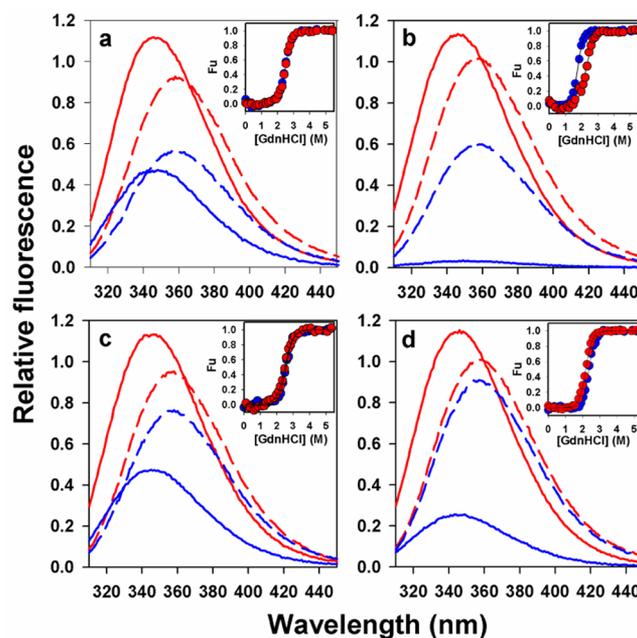
**Tryptophan Fluorescence Changes Due to FRET.** Four different single-tryptophan- and single-cysteine-containing variants of MNEI were studied by FRET. In each, the single tryptophan, W4, was the FRET donor (D) and a thionitrobenzoate (TNB) moiety, attached to a differently located thiol (of C29, C42, C53, or C97), was the FRET acceptor (A) (Figure 1).



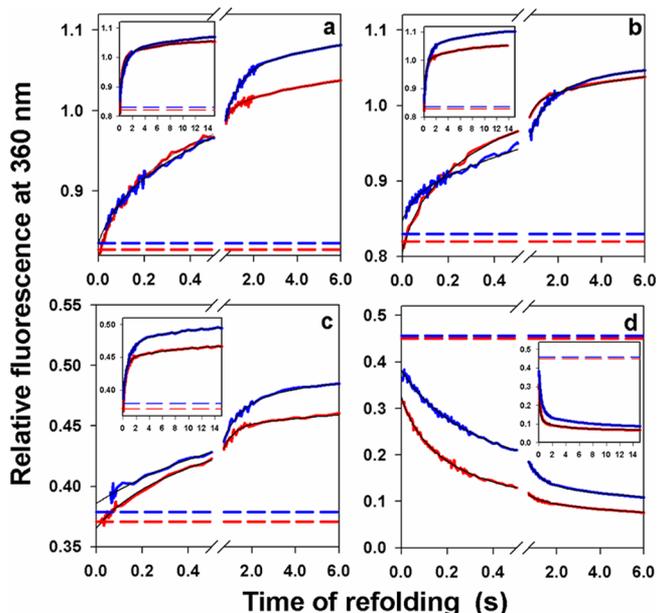
**Figure 1.** Structure of MNEI. The figure shows the positions of the residues that have been mutated to cysteine in the different single-cysteine-containing mutant forms of MNEI used for the multisite FRET measurements. The sole helix is formed by the sequence stretch of G10–G28. The single tryptophan residue (W4) used as the donor fluorophore is also shown. The structure was drawn from Protein Data Bank entry 1FA3 using PyMOL (version 1.30, Schrödinger, LLC).

The fluorescence emission spectra of the unlabeled mutant variants Cys29, Cys42, Cys53, and Cys97 and their labeled counterparts, Cys29-TNB, Cys42-TNB, Cys53-TNB, and Cys97-TNB, respectively, in their folded and unfolded states, are shown in Figure 2. For all mutant variants, the fluorescence intensity of the labeled protein was lower than that of the corresponding unlabeled counterpart, and the fluorescence intensity was lower at all the wavelengths, which is a hallmark of FRET. The extent of fluorescence quenching was greater in the native protein than in the unfolded protein, because the D–A distance is shorter in the native protein. The extent of quenching was different for the different mutant variants, as expected for cases in which quenching is distance-dependent, and due to FRET.<sup>53</sup>

**Labeling Does Not Affect the Stability or Folding Pathway of MNEI.** The stabilities of the TNB-labeled and unlabeled variants of the different single-cysteine-containing variants of MNEI are similar (Figure 2, inset). The stability of Cys42-TNB is 0.8 kcal/mol lower than that of the unlabeled form (Figure 2b, inset). Because C42 is completely buried in the native state of Cys42, the modification of this residue, by the addition of the TNB group, may have disrupted packing interactions and caused the decrease in stability. The kinetic and equilibrium experiments were performed by monitoring the tryptophan fluorescence emission at 360 nm, which is close to the isoemissive point of the fluorescence spectra of the unfolded and native forms of the unlabeled protein (Figure 2). Most of the change in the fluorescence signal in the case of the labeled protein (donor in the presence of acceptor) can then be attributed to the change in distance, because the fluorescence of the unlabeled (donor only) protein changes only marginally (Figure 3a,b). The refolding of MNEI proceeds in three kinetic phases when it is monitored by fluorescence (Figure 3), as

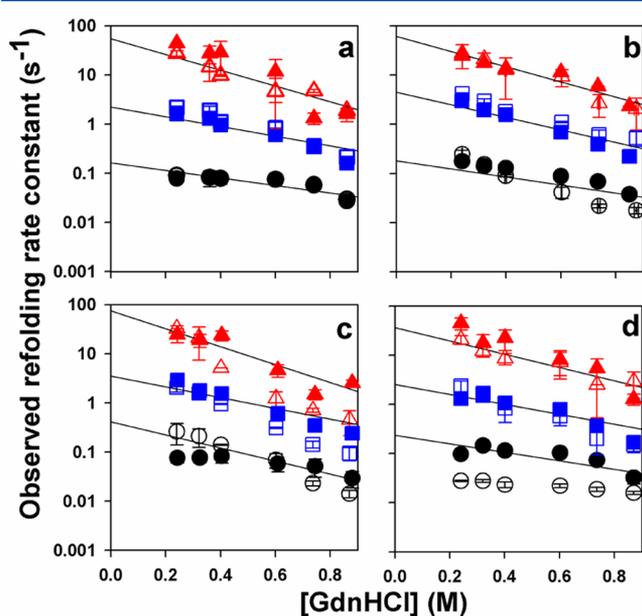


**Figure 2.** Fluorescence emission spectra of labeled (blue) and unlabeled (red) single-cysteine-containing mutant forms of MNEI: (a) Cys29 and Cys29-TNB, (b) Cys42 and Cys42-TNB, (c) Cys53 and Cys53-TNB, and (d) Cys97 and Cys97-TNB. The spectra of the native protein are shown as solid lines, and those of the unfolded protein in 4 M GdnHCl are shown as dashed lines. The spectra were normalized to the fluorescence value of the native unlabeled protein at 360 nm. The equilibrium unfolding transitions of the labeled (blue) and unlabeled (red) protein variants are shown in the inset.



**Figure 3.** Representative refolding kinetic traces of the labeled and unlabeled forms of MNEI. Panels a and c show the refolding traces of Cys29 and Cys29-TNB, respectively, and panels b and d show the refolding traces of Cys42 and Cys42-TNB, respectively. The red and blue lines represent the refolding traces obtained at 0.24 and 0.4 M GdnHCl, respectively. The extrapolated unfolded protein signals at 0.24 and 0.4 M GdnHCl are shown as the dashed red and blue lines, respectively. The inset in each panel shows the complete kinetic curves. The black lines through the data are three-exponential fits to the data.

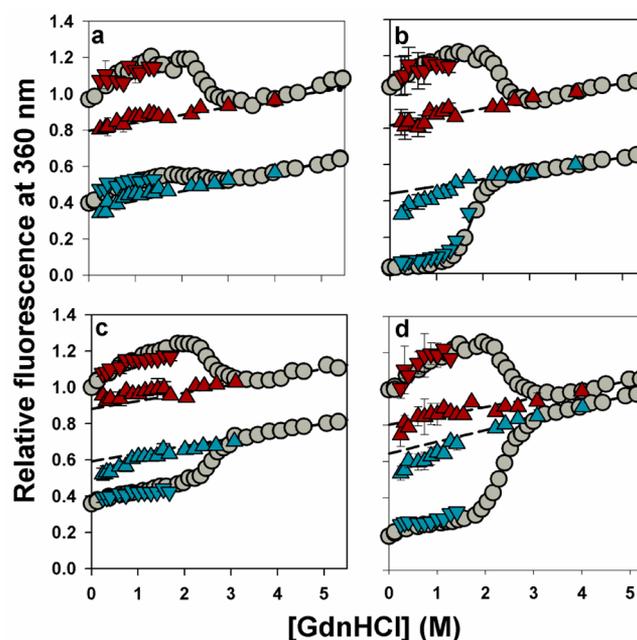
described previously.<sup>58</sup> The fluorescence signal of Cys29-TNB shows an unusual increase with folding time (Figure 3c), whereas the fluorescence signal of Cys42-TNB shows the expected decrease (Figure 3d). The rate constants of the unlabeled and the labeled variants were found to be similar (Figure 4) except for that of the slow phase of Cys97-TNB (Figure 4d).



**Figure 4.** Comparison of the rate constants of the three refolding phases of MNEI. Refolding rate constants of the labeled (empty symbols) and unlabeled (filled symbols) variants of (a) Cys29, (b) Cys42, (c) Cys53, and (d) Cys97: (triangles) very fast phase, (squares) fast phase, and (circles) slow phase of refolding. The solid lines through the data points were drawn to guide the eye. The error bars represent the standard errors obtained from two independent experiments.

**MNEI Undergoes Denaturant-Induced Expansion in the Native and Unfolded States.** The equilibrium unfolding transitions of both the labeled and unlabeled proteins are shown in Figure 5. The signal for the unlabeled protein changes only marginally because the signal is collected at 360 nm, close to the isoemissive point of the native and unfolded forms of the unlabeled protein variants. The magnitude of the signal of the labeled protein increases in a sigmoidal manner with an increase in GdnHCl concentration (Figure 5). The increase in denaturant concentration causes the unfolding of the proteins, which results in an increase in the D–A distance, resulting in an increase in tryptophan fluorescence. The extent of the increase in the fluorescence of labeled protein is different for different segments of the protein; this indicates that the increase is due to FRET.<sup>53</sup> MNEI undergoes polymer-like expansion when the denaturant concentration is increased in the native and unfolded protein baseline regions, which is evident from the increase in the fluorescence of the labeled protein with an increase in GdnHCl concentration, for all four distances monitored (Figure 5).

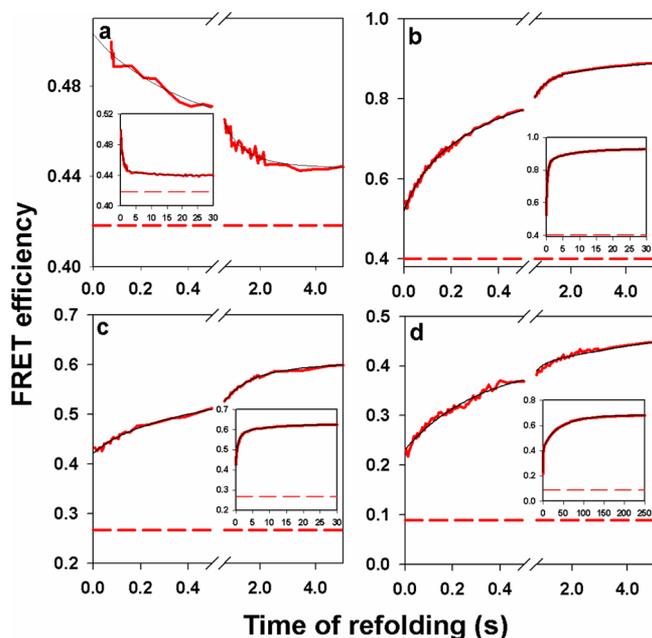
**MNEI Undergoes Chain Collapse within 2 ms of Folding.** Figure 5 compares the amplitude of the signal change observed in the kinetic refolding experiment to that observed in the equilibrium unfolding transition. The  $t = 0$  signal in the case of the unlabeled proteins falls on the extrapolated unfolded



**Figure 5.** Comparison of the kinetic and equilibrium amplitudes of fluorescence change for the four single-cysteine-containing variants of MNEI. Data for (a) Cys29 and Cys29-TNB, (b) Cys42 and Cys42-TNB, (c) Cys53 and Cys53-TNB, and (d) Cys97 and Cys97-TNB: (gray circles) equilibrium unfolding transitions of the labeled (bottom) and unlabeled (top) proteins, (blue triangles)  $t = 0$  and  $t = \infty$  signals of the kinetic refolding traces of the labeled protein, and (red triangles) extrapolated  $t = 0$  values and  $t = \infty$  values of the kinetic traces of folding of the unlabeled protein. In each panel, the dashed black line represents the unfolded protein baseline extrapolated to lower GdnHCl concentrations.

protein baseline (Figure 5). The burst phase loss of the fluorescence signal in the case of the labeled proteins shows that the distances have contracted within the dead time of 2 ms. For Cys29-TNB, the kinetic  $t = 0$  points fall on the extrapolated unfolded protein baseline. This indicates that for the species populated within the 2 ms dead time of stopped flow mixing, the W4–C29TNB distance has contracted to the value expected for the unfolded state under refolding conditions. This distance appears to have undergone nonspecific contraction (Figure 5a). For Cys42-TNB, Cys53-TNB, and Cys97-TNB, the kinetic  $t = 0$  points fall below the extrapolated unfolded baseline, indicating that the contraction of the W4–C42TNB, W4–C53TNB, and W4–C97TNB distances is greater than that expected by only nonspecific contraction induced by solvent change. The contraction of these distances would appear to be associated with the formation of specific structural elements in the protein at 2 ms of folding (Figure 5b–d). At higher (>1 M) GdnHCl concentrations, however, all the distances appear to undergo only a nonspecific solvent-induced compaction (Figure 5).

The kinetic traces monitored using FRET, calculated from the fluorescence signals of the labeled ( $F_{DA}$ ) and unlabeled ( $F_D$ ) proteins using eq 1, show an initial burst phase increase from the value seen for the unfolded state, for all four distances probed (Figure 6). This initial increase in FRET efficiency is indicative of a collapse transition occurring within 2 ms. The W4–C42, W4–C53, and W4–C97 segments show further increases in FRET efficiency as refolding proceeds, suggesting that further compaction occurs along with structure formation

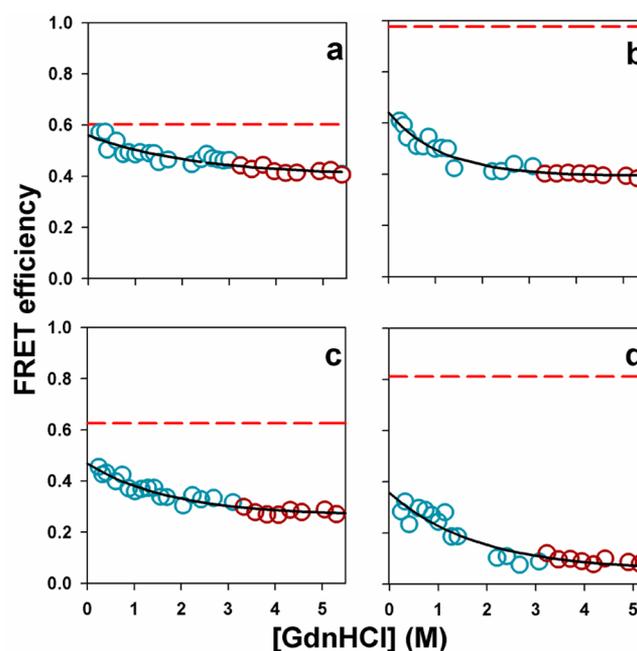


**Figure 6.** FRET efficiency-monitored refolding traces of (a) W4–C29, (b) W4–C42, (c) W4–C53, and (d) W4–C97. The FRET efficiencies were calculated from the kinetic traces of refolding of the labeled and unlabeled proteins at 0.4 M GdnHCl. The very fast and fast refolding phases are shown in the main panels, and the complete kinetic curves are shown in the insets. The black line through the data is a three-exponential fit to the data. The FRET efficiency in the unfolded state in 4 M GdnHCl is shown as a dashed red line.

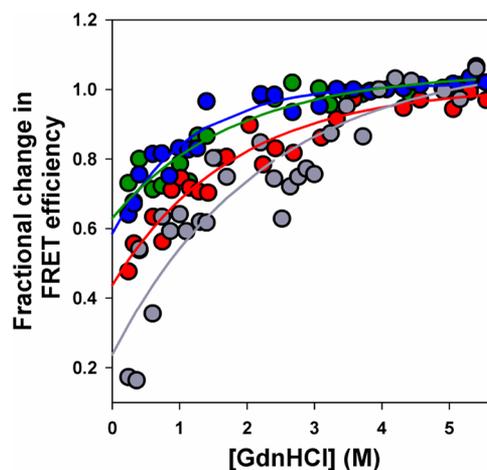
(Figure 6b–d). Rather unusually, the W4–C29 segment shows a decrease in FRET efficiency as the folding reaction proceeds, which indicates that it expands with increasing folding time (Figure 6a).

**Asynchronous Compaction of Different Segments of MNEI.** From the kinetic  $t = 0$  signals obtained from the kinetic traces of folding of the labeled and unlabeled proteins (Figure 5), the FRET efficiency of the collapsed state (kinetic MG) formed at 2 ms was calculated using eq 1. Figure 7 shows that the FRET efficiency of the collapsed species decreases with an increase in GdnHCl concentration in an apparently exponential manner, and its dependence on GdnHCl concentration continues into that of the unfolded state. This indicates that the process that results in the formation of the kinetic MG is noncooperative. It was further observed that the dependence on GdnHCl concentration of the fractional change in FRET efficiency calculated using eq 2 is different for the four distances probed (Figure 8), indicating that the four segments have contracted asynchronously, and that the four segments of structure in the kinetic MG have different stabilities. The midpoints for the denaturant dependences of the fractional change in FRET efficiency are different for the four distances studied (Figure 8). The W4–C29 segment, which ultimately forms the helix, has contracted most in the kinetic MG (Figure 8).

**The Helix of MNEI Is Absent in the Kinetic MG.** To examine the secondary structural content of the kinetic MG, the refolding reaction was monitored by measurement of the change in ellipticity at 222 nm, which probes the formation of the single  $\alpha$ -helix in MNEI. The extrapolated  $t = 0$  values of the kinetic folding traces monitored by far-UV CD at 222 nm fall on the extrapolated unfolded protein baseline (Figure 9b). This

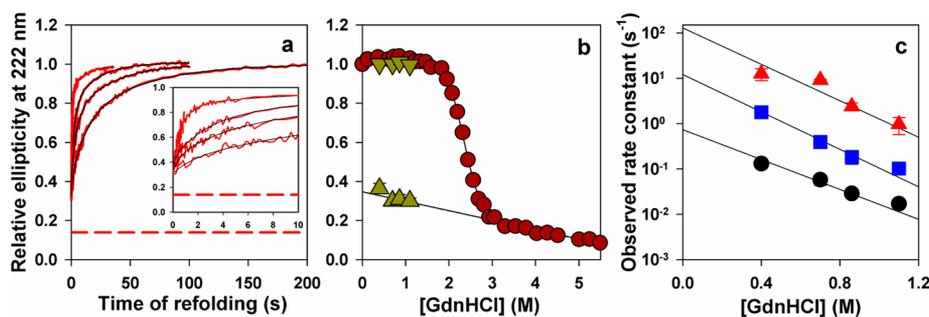


**Figure 7.** FRET efficiency of the initially collapsed form at different GdnHCl concentrations: (a) W4–C29, (b) W4–C42, (c) W4–C53, and (d) W4–C97. The cyan circles represent the FRET efficiencies of the initially collapsed form (kinetic MG), which were calculated from the extrapolated  $t = 0$  values of the kinetic traces of folding of the labeled and unlabeled proteins. The dark red circles were calculated from the equilibrium unfolding curves of the labeled and unlabeled proteins. The black lines through the data points are single-exponential fits to the data. In each panel, the FRET efficiency of the native state is shown as a dashed red line.

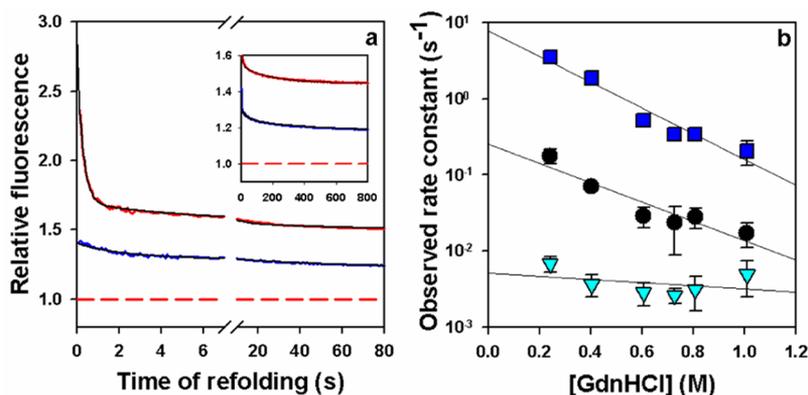


**Figure 8.** Fractional change in FRET efficiency of the initially collapsed form. The fractional changes in FRET efficiency of the initially collapsed form at different GdnHCl concentrations are shown for segments W4–C29 (gray), W4–C42 (blue), W4–C53 (red), and W4–C97 (green). The solid lines through the data points have been drawn to guide the eye.

shows that the helix is absent in the kinetic MG. Helical structure forms in three kinetic phases whose rate constants (Figure 9a,c) are similar to those monitored by fluorescence (Figure 4). The sole helix in MNEI therefore forms in multiple kinetic phases. The W4-C29TNB distance monitors the segment of the protein that folds to form the  $\alpha$ -helix and is therefore an alternative probe for the formation of helical



**Figure 9.** Refolding of MNEI monitored by far-UV CD. (a) Representative kinetic traces of the folding of Cys42 (wt MNEI) at different GdnHCl concentrations from 0.4 to 1.0 M. The solid red lines represent the data points, and the black lines represent multiexponential fits through the data. The signal of the unfolded protein in 4 M GdnHCl is shown as a dashed red line. (b) Comparison of the equilibrium and kinetic amplitudes of the far-UV CD signal change: (green upward-pointing triangles) extrapolated  $t = 0$  signals from the fits to the kinetic traces and (green downward-pointing triangles)  $t = \infty$  signals of the kinetic traces. The solid black line is the extrapolated unfolded protein baseline. (c) Refolding rate constants obtained from the fits to the kinetic traces: (red triangles) very fast phase, (blue squares) fast phase, and (black circles) slow phase of refolding. The signals have been normalized to a value of 1 for the signal of the native protein in 0 M GdnHCl. The error bars represent the standard errors obtained from two independent experiments. The solid black lines through the data points are linear regression fits.



**Figure 10.** ANS-monitored refolding kinetics of MNEI. (a) Representative refolding kinetic traces of Cys42 at 0.24 M GdnHCl (red line) and 0.6 M GdnHCl (blue line). The fast and slow refolding phases are shown. The very slow phase is shown in the inset. The signals were normalized to the signal of the unfolded protein in 4 M GdnHCl, which is shown as a dashed red line. (b) Rate constants of the different refolding phases of MNEI observed at different GdnHCl concentrations: (blue squares) fast phase, (black circles) slow phase, and (cyan triangles) very slow phase. The black lines through the data are linear regression fits.

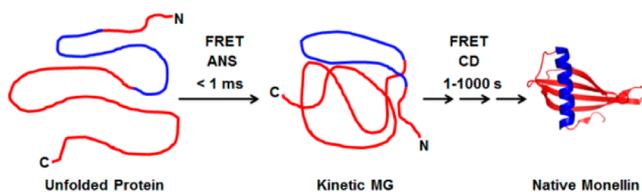
structure. The rates at which the W4-C29TNB distance increases (Figure 4a) are similar to the rates at which the magnitude of the far-UV CD signal at 222 nm increases (Figure 9c).

**The Kinetic MG of MNEI Has Exposed Hydrophobic Patches.** ANS fluorescence-monitored refolding experiments conducted using stopped flow mixing show a substantial increase in ANS fluorescence consequent to its binding, which occurs in the burst phase (Figure 10a). This indicates that the kinetic MG of MNEI has solvent-accessible hydrophobic patches that can bind to ANS, which are absent in the unfolded state and buried in the native state. As folding proceeds further, ANS is expelled out of the core, which becomes dry, resulting in a decrease in the fluorescence of ANS (Figure 10a). The decrease happens in three kinetic phases (Figure 10b), two of which have rates similar to those of the fast and slow phases of fluorescence-monitored refolding. The very fast phase observed using tryptophan fluorescence, FRET, and far-UV CD as probes is silent to ANS fluorescence. An additional very slow phase, which is silent to tryptophan fluorescence, FRET, and far-UV CD, is observed with ANS fluorescence. The lack of denaturant dependence as well as the time scale of this very slow phase suggests that it may be

associated with a late proline isomerization step during folding (Figure 10b).<sup>58</sup>

■ DISCUSSION

The fast time scales of early folding events make it difficult to delineate the temporal order of transitions such as chain compaction and secondary structure formation. In this study, the formation of secondary structure was probed in a slow folding protein MNEI, which is known to undergo chain collapse within 1 ms of refolding (Figure 11).<sup>57</sup> The size of the



**Figure 11.** Cartoon diagram showing the collapse of the protein within the stopped flow mixing dead time of 2 ms. The helical segment is colored blue. The helix is not present in the collapsed form and forms on the time scale of 1–1000 s. The probes used for detecting the different stages of folding are shown above the arrow.

initially collapsed species (kinetic MG) formed at a few milliseconds of refolding was monitored by multisite FRET, which showed that the initial compaction occurs noncooperatively and leads to the formation of some specific structure in the kinetic MG. The kinetic MG is, however, shown to lack helical secondary structure that forms subsequently in slower kinetic phases.

**The Initial Collapse Reaction of MNEI Is a Continuous Transition.** FRET efficiencies measured in different segments of the kinetic MG populated at 2 ms of the refolding reaction show exponential dependences on the concentration of the denaturant, which merge into those of the FRET efficiencies measured in the unfolded state at higher denaturant concentrations (Figure 7). If the compaction event happening within the first 2 ms of folding were a barrier-limited transition, the product of such a reaction would have exhibited a sigmoidal dependence of the FRET efficiency on the denaturant concentration. It can thus be concluded that the initial chain compaction of MNEI is a noncooperative barrier-less transition. Previously, ensemble multisite FRET measurements had shown that initial chain collapses in the case of barstar<sup>21,22,50,52,61</sup> and the SH3 domain of PI3 kinase<sup>41</sup> are also gradual transitions. Single-molecule FRET measurements have also shown that the compaction of the polypeptide chain at low denaturant concentrations is a barrier-less gradual transition in the case of cold shock protein CspTm,<sup>62</sup> chymotrypsin inhibitor 2 (CI2),<sup>63</sup> RNase H,<sup>64</sup> and protein L.<sup>65</sup> On the other hand, collapse was suggested to be a barrier-limited transition in the case of cytochrome *c*, based on the observation of exponential folding kinetics with an Arrhenius-like temperature dependence.<sup>66</sup> Exponential kinetics may not, however, be a signature of a barrier-limited transition; indeed, it was shown that a continuum of states with stabilities linearly dependent on denaturant concentration can also show exponential kinetics.<sup>67,68</sup>

The native and unfolded protein baselines of the denaturant-induced equilibrium unfolding curves for all the labeled mutant forms of MNEI have a significant dependence on denaturant concentration (Figure 5). Such baselines, when observed with FRET, are indicative of changes in the dimensions of the protein upon addition of denaturant.<sup>41,46,50,53,64</sup> A recent study based on molecular dynamics simulations and FCS measurements has suggested that side chain-mediated interactions are responsible for the expansion of the polypeptide chain in solutions containing denaturants.<sup>69</sup> The observation that protein chains expand with an increase in denaturant concentration indicates that the evolutionarily designed polypeptide chain behaves as a simple homopolymer when the solvent conditions are altered. In fact, the unfolding of MNEI, probed by time-resolved FRET, could be explained by a Rouse-like model of polymer physics, with a few intrachain interactions.<sup>54</sup>

**The Kinetic MG State of MNEI Is Highly Heterogeneous.** The four segments probed by multisite FRET behave differently when the solvent conditions are changed to initiate folding. The distance monitored in Cys29-TNB undergoes a nonspecific contraction (Figure 5a), whereas the other three distances appear to undergo a specific collapse transition resulting in the contraction of the segments being greater than that expected from extrapolation of the unfolded protein baseline (Figure 5b–d). Moreover, the distance monitored in Cys29-TNB, unlike the other distances probed, was observed to have contracted, in the kinetic MG formed at the lower

GdnHCl concentrations used, to a distance that was shorter than that in the native state. This result is perhaps not unexpected for a segment that undergoes nonspecific collapse (Figure 5a).

The heterogeneity of structure in the kinetic MG is evident from the asynchronous contraction of different segments of the protein upon a change in denaturant concentration (Figure 8). The heterogeneous nature of the kinetic MG at a few milliseconds of folding has also been observed for barstar,<sup>50,52</sup> the  $\alpha$  subunit of tryptophan synthase,<sup>70</sup> and staphylococcal nuclease.<sup>71</sup> The heterogeneity of the collapsed species can be modulated by a change in experimental conditions. For example, the fluorescence and CD properties of the burst phase intermediate formed initially during the folding of barstar could be modulated by denaturant, salt, and osmolytes.<sup>16–18</sup> In the case of apomyoglobin, the helical content of the collapsed intermediate formed at a few milliseconds of folding was found to increase with mutations that stabilized individual helices.<sup>19</sup>

It is likely that the kinetic MG of monellin is stabilized by non-native hydrophobic interactions. Such interactions are likely to transiently stabilize the sequence segment that has undergone nonspecific collapse in the kinetic MG and goes on to form the helix (see above). Transient non-native hydrophobic interactions have been implicated in the stabilization of folding intermediates.<sup>12,72–75</sup> However, at present, there is no direct evidence that such interactions play a role either in stabilizing the kinetic MG of monellin or in modulating its subsequent folding to the N state, as has been seen for other proteins.<sup>72,76–79</sup>

**The Sole Helix of MNEI Is Absent in the Kinetic MG at 2 ms of Folding.** It is difficult to estimate the secondary structural content of kinetic MGs by CD because far-UV CD signals can change because of burial of aromatic amino acids.<sup>42,43</sup> In fact, the far-UV CD signal of the kinetic MG can be mimicked by nonfoldable analogues of proteins.<sup>80,81</sup> In the study presented here, to delineate the status of the helix in the kinetic MG of MNEI, FRET was used to follow the contraction and expansion of the sequence segment that forms the helix (W4–C29) (Figure 1). The W4–C29 segment collapses in the kinetic MG to a distance that is shorter than the distance in the native state, which is not surprising for a nonspecific, solvent-induced compaction (Figure 5a). This observation clearly shows that the kinetic MG of MNEI is devoid of the sole helix. This result was further confirmed by far-UV CD measurements (Figure 9). As the helix subsequently forms from the kinetic MG, the D–A distance in W4–C29 segment expands to accommodate the helix (Figure 6a). This result illustrates the usefulness of FRET in establishing the secondary structural content of kinetic MGs. A similar FRET strategy was used to delineate the secondary structural content of the earliest intermediate populated during the folding of adenylate kinase, where it was shown that fast chain collapse precedes the formation of secondary structural elements.<sup>24</sup>

The results of the FRET (Figure 6a) and far-UV CD (Figure 9) experiments clearly show that the kinetic MG of MNEI lacks helical secondary structure. It supports the view that the compaction of the polypeptide chain is responsible for the formation of secondary structure. Theoretical studies have shown that any polymer with attractive forces within the chain can undergo compaction, and as a result, they form secondary structural elements.<sup>82</sup> It has been suggested<sup>30</sup> that the major driving forces for the formation of an  $\alpha$ -helix are the hydrophobic effect and the van der Waals forces in the closely

packed helix. Helix formation was suggested to be driven by hydrophobic collapse, and hydrogen bonding was suggested to be a consequence of the structural constraint induced by the energetically unfavorable burial of hydrogen bonding groups.<sup>30</sup> Helices in isolation are not stable. A classic example is the myoglobin helices that are unstable in isolation, showing that the hydrophobic effect and packing interactions play an important role in promoting secondary structure formation.<sup>83,84</sup>

It is also known that packing interactions between the side chains of amino acid residues are a major stabilizing force in isolated helices.<sup>29</sup> Not surprisingly, the chain B subunit of double-chain monellin, which contains the residue stretch that folds to form the helix, has no helicity in isolation.<sup>85</sup>

**The Helix of MNEI Forms as the Core Becomes Compact.** The kinetic MG of MNEI at 2 ms of folding is also capable of binding to the hydrophobic dye ANS (Figure 10), which indicates that the core of the protein is not yet consolidated. The helix rises from this less compact (compared to the native state) kinetic MG, as observed by FRET measurement of the W4–C29 (Figure 6a) distance and by far-UV CD measurement (Figure 9), with a concomitant decrease in ANS fluorescence (Figure 10a). The hydrophobic environment created when water is expelled from the core after initial chain collapse has occurred is likely to facilitate the formation of hydrogen bonding between the backbone amides, resulting in the formation of the helix.

Recent site-specific thiol labeling-monitored unfolding studies under native conditions have shown that C13 in the helix unfolds first, and that the packing interactions in the  $\beta$ -sheet are the last ones to dissolve.<sup>86</sup> In the refolding study presented here, it is observed that under stabilizing (nativelike) conditions, specific structure has formed within 2 ms in the segments that constitute the  $\beta$ -sheet, namely, W4–C42, W4–C53, and W4–C97 (Figure 5b–d). The segment that constitutes the helix (W4–C29) undergoes a nonspecific solvent-induced compaction (Figure 5a), and the helix forms from the asynchronously collapsed kinetic MG in three kinetic phases, as observed by FRET and far-UV CD measurements. These observations suggest that the sequence of folding events observed in the kinetic refolding experiments is in the reverse order of the sequence of unfolding events observed in the native state thiol labeling measurements.

## CONCLUSION

In this study, the kinetic MG of MNEI is shown, by multisite FRET measurements, to be compact at 2 ms of folding. The initial chain collapse reaction is shown to be noncooperative and gradual from the observed asynchronous compaction of different segments of the polypeptide chain. The kinetic MG is devoid of any helical secondary structure as observed by far-UV CD, as well as by FRET between the D–A pair placed suitably in the sequence segment that goes on to form the helix. The sole helix of MNEI was shown to form in three kinetic phases as monitored by FRET and far-UV CD measurements. The helix rises from the asynchronously collapsed kinetic MG as the core becomes compact. It appears that the formation of backbone hydrogen bonds of the helix might be facilitated by the expulsion of water from the compact kinetic MG. This study illustrates the importance of initial chain collapse and packing interactions in stabilizing the secondary structural components of proteins.

## AUTHOR INFORMATION

### Corresponding Author

\*Fax: 91-80-23636662. E-mail: jayant@ncbs.res.in.

### Funding

This work was funded by the Tata Institute of Fundamental Research and by the Department of Science and Technology, Government of India. J.B.U. is a recipient of a J. C. Bose National Fellowship from the Government of India.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank members of our laboratory for discussions and comments on the manuscript.

## ABBREVIATIONS

MNEI, single-chain monellin; MG, molten globule; GdnHCl, guanidine hydrochloride; FRET, fluorescence (Forster) resonance energy transfer; CD, circular dichroism; ANS, 1-anilinonaphthalene 8-sulfonate; TNB, thionitrobenzoate.

## REFERENCES

- (1) Parr, G. R., and Taniuchi, H. (1983) Evidence of a compact structure for kinetic intermediates in the folding of a fragment complex of tuna cytochrome c. *J. Biol. Chem.* 258, 3759–3763.
- (2) Kuwajima, K. (1989) The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins: Struct., Funct., Genet.* 6, 87–103.
- (3) Fink, A. L. (1995) Compact Intermediate States in Protein Folding. *Annu. Rev. Biophys. Biomol. Struct.* 24, 495–522.
- (4) Ptitsyn, O. B. (1995) How the molten globule became. *Trends Biochem. Sci.* 20, 376–379.
- (5) Baldwin, R. L., and Rose, G. D. (2013) Molten globules, entropy-driven conformational change and protein folding. *Curr. Opin. Struct. Biol.* 23, 4–10.
- (6) Kuwajima, K., Semisotnov, G. V., Finkelstein, A. V., Sugai, S., and Ptitsyn, O. B. (1993) Secondary structure of globular proteins at the early and the final stages in protein folding. *FEBS Lett.* 334, 265–268.
- (7) Udgaonkar, J. B., and Baldwin, R. L. (1990) Early folding intermediate of ribonuclease A. *Proc. Natl. Acad. Sci. U. S. A.* 87, 8197–8201.
- (8) Jennings, P. A., and Wright, P. E. (1993) Formation of a Molten Globule Intermediate Early in the Kinetic Folding Pathway of Apomyoglobin. *Science (Washington, DC, U. S.)* 262, 892–896.
- (9) Hamada, D., Segawa, S., and Goto, Y. (1996) Non-native [ $\alpha$ ]-helical intermediate in the refolding of [ $\beta$ ]-lactoglobulin, a predominantly [ $\beta$ ]-sheet protein. *Nat. Struct. Biol.* 3, 868–873.
- (10) Capaldi, A. P., Kleantous, C., and Radford, S. E. (2002) Im7 folding mechanism: misfolding on a path to the native state. *Nat. Struct. Biol.* 9, 209–216.
- (11) Li, J., Shinjo, M., Matsumura, Y., Morita, M., Baker, D., Ikeguchi, M., and Kihara, H. (2007) An  $\alpha$ -Helical Burst in the src SH3 Folding Pathway. *Biochemistry* 46, 5072–5082.
- (12) Dasgupta, A., and Udgaonkar, J. B. (2012) Transient Non-Native Burial of a Trp Residue Occurs Initially during the Unfolding of a SH3 Domain. *Biochemistry* 51, 8226–8234.
- (13) Matsumura, Y., Shinjo, M., Kim, S. J., Okishio, N., Gruebele, M., and Kihara, H. (2013) Transient Helical Structure during PI3K and Fyn SH3 Domain Folding. *J. Phys. Chem. B* 117, 4836–4843.
- (14) Peng, Z. Y., and Kim, P. S. (1994) A protein dissection study of a molten globule. *Biochemistry* 33, 2136–2141.
- (15) Uversky, V. N., and Fink, A. L. (1999) Do protein molecules have a native-like topology in the pre-molten globule state? *Biochemistry (Moscow)* 64, 552–555.

- (16) Agashe, V. R., Shastry, M. C. R., and Udgaonkar, J. B. (1995) Initial hydrophobic collapse in the folding of barstar. *Nature* 377, 754–757.
- (17) Pradeep, L., and Udgaonkar, J. B. (2002) Differential Salt-induced Stabilization of Structure in the Initial Folding Intermediate Ensemble of Barstar. *J. Mol. Biol.* 324, 331–347.
- (18) Pradeep, L., and Udgaonkar, J. B. (2004) Osmolytes induce structure in an early intermediate on the folding pathway of barstar. *J. Biol. Chem.* 279, 40303–40313.
- (19) Kiefhaber, T., and Baldwin, R. L. (1995) Intrinsic Stability of Individual  $\alpha$  Helices Modulates Structure and Stability of the Apomyoglobin Molten Globule Form. *J. Mol. Biol.* 252, 122–132.
- (20) Khorasanizadeh, S., Peters, I. D., and Roder, H. (1996) Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. *Nat. Struct. Biol.* 3, 193–205.
- (21) Rami, B. R., and Udgaonkar, J. B. (2002) Mechanism of Formation of a Productive Molten Globule Form of Barstar†. *Biochemistry* 41, 1710–1716.
- (22) Rami, B. R., Krishnamoorthy, G., and Udgaonkar, J. B. (2003) Dynamics of the Core Tryptophan during the Formation of a Productive Molten Globule Intermediate of Barstar†. *Biochemistry* 42, 7986–8000.
- (23) Magg, C., and Schmid, F. X. (2004) Rapid Collapse Precedes the Fast Two-state Folding of the Cold Shock Protein. *J. Mol. Biol.* 335, 1309–1323.
- (24) Ratner, V., Amir, D., Kahana, E., and Haas, E. (2005) Fast Collapse but Slow Formation of Secondary Structure Elements in the Refolding Transition of E.coli Adenylate Kinase. *J. Mol. Biol.* 352, 683–699.
- (25) Udgaonkar, J. B. (2013) Polypeptide chain collapse and protein folding. *Arch. Biochem. Biophys.* 531, 24–33.
- (26) Sadqi, M., Lapidus, L. J., and Munoz, V. (2003) How fast is protein hydrophobic collapse? *Proc. Natl. Acad. Sci. U. S. A.* 100, 12117–12122.
- (27) Nettels, D., Gopich, I. V., Hoffmann, A., and Schuler, B. (2007) Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl. Acad. Sci. U. S. A.* 104, 2655–2660.
- (28) Kubelka, J., Hofrichter, J., and Eaton, W. A. (2004) The protein folding “speed limit”. *Curr. Opin. Struct. Biol.* 14, 76–88.
- (29) Padmanabhan, S., and Baldwin, R. L. (1994) Helix-stabilizing Interaction Between Tyrosine and Leucine or Valine when the Spacing is  $i, j + 4$ . *J. Mol. Biol.* 241, 706–713.
- (30) Yang, A., and Honig, B. (1995) Free Energy Determinants of Secondary Structure Formation: I.  $\alpha$ -Helices. *J. Mol. Biol.* 252, 351–365.
- (31) Akiyama, S., Takahashi, S., Ishimori, K., and Morishima, I. (2000) Stepwise formation of  $\alpha$ -helices during cytochrome c folding. *Nat. Struct. Biol.* 7, 514–520.
- (32) Uzawa, T., Akiyama, S., Kimura, T., Takahashi, S., Ishimori, K., Morishima, I., and Fujisawa, T. (2004) Collapse and search dynamics of apomyoglobin folding revealed by submillisecond observations of  $\alpha$ -helical content and compactness. *Proc. Natl. Acad. Sci. U. S. A.* 101, 1171–1176.
- (33) Hoffmann, A., Kane, A., Nettels, D., Hertzog, D. E., Baumgärtel, P., Lengfeld, J., Reichardt, G., Horsley, D. a, Seckler, R., Bakajin, O., and Schuler, B. (2007) Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* 104, 105–110.
- (34) Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L. J., Dobson, C. M., and Schwalbe, H. (2002) Long-range interactions within a nonnative protein. *Science* 295, 1719–1722.
- (35) Mok, K. H., Kuhn, L. T., Goez, M., Day, I. J., Lin, J. C., Andersen, N. H., and Hore, P. J. (2007) A pre-existing hydrophobic collapse in the unfolded state of an ultrafast folding protein. *Nature* 447, 106–109.
- (36) Hodsdon, M. E., and Frieden, C. (2001) Intestinal Fatty Acid Binding Protein: The Folding Mechanism As Determined by NMR Studies†. *Biochemistry* 40, 732–742.
- (37) Kazmirski, S. L., Wong, K.-B., Freund, S. M. V., Tan, Y.-J., Fersht, A. R., and Daggett, V. (2001) Protein folding from a highly disordered denatured state: The folding pathway of chymotrypsin inhibitor 2 at atomic resolution. *Proc. Natl. Acad. Sci. U. S. A.* 98, 4349–4354.
- (38) Lindorff-Larsen, K., Piana, S., Dror, R. O., and Shaw, D. E. (2011) How Fast-Folding Proteins Fold. *Science* 334, 517–520.
- (39) Arai, M., Kondrashkina, E., Kayatekin, C., Matthews, C. R., Iwakura, M., and Bilsel, O. (2007) Microsecond Hydrophobic Collapse in the Folding of Escherichia coli Dihydrofolate Reductase, an  $\alpha/\beta$ -Type Protein. *J. Mol. Biol.* 368, 219–229.
- (40) Welker, E., Maki, K., Shastry, M. C. R., Juminaga, D., Bhat, R., Scheraga, H. a, and Roder, H. (2004) Ultrarapid mixing experiments shed new light on the characteristics of the initial conformational ensemble during the folding of ribonuclease A. *Proc. Natl. Acad. Sci. U. S. A.* 101, 17681–17686.
- (41) Dasgupta, A., and Udgaonkar, J. B. (2010) Evidence for Initial Non-specific Polypeptide Chain Collapse During the Refolding of the SH3 Domain of PI3 Kinase. *J. Mol. Biol.* 403, 430–445.
- (42) Chakraborty, A., Kortemme, T., Padmanabhan, S., and Baldwin, R. L. (1993) Aromatic side-chain contribution to far-ultraviolet circular dichroism of helical peptides and its effect on measurement of helix propensities. *Biochemistry* 32, 5560–5565.
- (43) Vuilleumier, S., Sancho, J., Loewenthal, R., and Fersht, A. R. (1993) Circular dichroism studies of barnase and its mutants: Characterization of the contribution of aromatic side chains. *Biochemistry* 32, 10303–10313.
- (44) Kjaergaard, M., Nørholm, A., Hendus-Altenburger, R., Pedersen, S. F., Poulsen, F. M., and Kragelund, B. B. (2010) Temperature-dependent structural changes in intrinsically disordered proteins: Formation of  $\alpha$ -helices or loss of polyproline II? *Protein Sci.* 19, 1555–1564.
- (45) Amir, D., and Haas, E. (1987) Estimation of intramolecular distance distributions in bovine pancreatic trypsin inhibitor by site-specific labeling and nonradiative excitation energy-transfer measurements. *Biochemistry* 26, 2162–2175.
- (46) Lakshmikanth, G. S., Sridevi, K., Krishnamoorthy, G., and Udgaonkar, J. B. (2001) Structure is lost incrementally during the unfolding of barstar. *Nat. Struct. Biol.* 8, 799–804.
- (47) Lee, J. C., Engman, K. C., Tezcan, F. A., Gray, H. B., and Winkler, J. R. (2002) Structural features of cytochrome c folding intermediates revealed by fluorescence energy-transfer kinetics. *Proc. Natl. Acad. Sci. U. S. A.* 99, 14778–14782.
- (48) Teilum, K., Maki, K., Kragelund, B. B., Poulsen, F. M., and Roder, H. (2002) Early kinetic intermediate in the folding of acyl-CoA binding protein detected by fluorescence labeling and ultrarapid mixing. *Proc. Natl. Acad. Sci. U. S. A.* 99, 9807–9812.
- (49) Sridevi, K., and Udgaonkar, J. B. (2003) Surface Expansion Is Independent of and Occurs Faster than Core Solvation during the Unfolding of Barstar†. *Biochemistry* 42, 1551–1563.
- (50) Sinha, K. K., and Udgaonkar, J. B. (2005) Dependence of the size of the initially collapsed form during the refolding of barstar on denaturant concentration: Evidence for a continuous transition. *J. Mol. Biol.* 353, 704–718.
- (51) Magg, C., Kubelka, J., Holtermann, G., Haas, E., and Schmid, F. X. (2006) Specificity of the initial collapse in the folding of the cold shock protein. *J. Mol. Biol.* 360, 1067–1080.
- (52) Sinha, K. K., and Udgaonkar, J. B. (2007) Dissecting the Non-specific and Specific Components of the Initial Folding Reaction of Barstar by Multi-site FRET Measurements. *J. Mol. Biol.* 370, 385–405.
- (53) Jha, S. K., and Udgaonkar, J. B. (2009) Direct evidence for a dry molten globule intermediate during the unfolding of a small protein. *Proc. Natl. Acad. Sci. U. S. A.* 106, 12289–12294.
- (54) Jha, S. K., Dhar, D., Krishnamoorthy, G., and Udgaonkar, J. B. (2009) Continuous dissolution of structure during the unfolding of a small protein. *Proc. Natl. Acad. Sci. U. S. A.* 106, 11113–11118.

- (55) Orevi, T., Ben Ishay, E., Gershanov, S. L., Dalak, M., Amir, D., and Haas, E. (2014) Fast Closure of N-Terminal Long Loops but Slow Formation of  $\beta$  Strands Precedes the Folding Transition State of *Escherichia coli* Adenylate Kinase. *Biochemistry* 53, 3169–3178.
- (56) Ogata, C., Hatada, M., Tomlinson, G., Shin, W.-C., and Kim, S.-H. (1987) Crystal structure of the intensely sweet protein monellin. *Nature* 328, 739–742.
- (57) Kimura, T., Uzawa, T., Ishimori, K., Morishima, I., Takahashi, S., Konno, T., Akiyama, S., and Fujisawa, T. (2005) Specific collapse followed by slow hydrogen-bond formation of  $\beta$ -sheet in the folding of single-chain monellin. *Proc. Natl. Acad. Sci. U. S. A.* 102, 2748–2753.
- (58) Patra, A. K., and Udgaonkar, J. B. (2007) Characterization of the folding and unfolding reactions of single-chain monellin: Evidence for multiple intermediates and competing pathways. *Biochemistry* 46, 11727–11743.
- (59) Jha, S. K., Dasgupta, A., Malhotra, P., and Udgaonkar, J. B. (2011) Identification of Multiple Folding Pathways of Monellin Using Pulsed Thiol Labeling and Mass Spectrometry. *Biochemistry* 50, 3062–3074.
- (60) Agashe, V. R., and Udgaonkar, J. B. (1995) Thermodynamics of denaturation of barstar: evidence for cold denaturation and evaluation of the interaction with guanidine hydrochloride. *Biochemistry* 34, 3286–3299.
- (61) Sinha, K. K., and Udgaonkar, J. B. (2008) Barrierless evolution of structure during the submillisecond refolding reaction of a small protein. *Proc. Natl. Acad. Sci. U. S. A.* 105, 7998–8003.
- (62) Schuler, B., Lipman, E. A., and Eaton, W. A. (2002) Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* 419, 743–747.
- (63) Hamadani, K. M., and Weiss, S. (2008) Nonequilibrium single molecule protein folding in a coaxial mixer. *Biophys. J.* 95, 352–365.
- (64) Kuzmenkina, E. V., Heyes, C. D., and Ulrich Nienhaus, G. (2006) Single-molecule FRET study of denaturant induced unfolding of RNase H. *J. Mol. Biol.* 357, 313–324.
- (65) Sherman, E., and Haran, G. (2006) Coil-globule transition in the denatured state of a small protein. *Proc. Natl. Acad. Sci. U. S. A.* 103, 11539–11543.
- (66) Shastry, M. C. R., and Roder, H. (1998) Evidence for barrier-limited protein folding kinetics on the microsecond time scale. *Nat. Struct. Biol.* 5 (5), 385–392.
- (67) Parker, M. J., and Marqusee, S. (1999) The cooperativity of burst phase reactions explored. *J. Mol. Biol.* 293, 1195–1210.
- (68) Hagen, S. J. (2003) Exponential decay kinetics in “downhill” protein folding. *Proteins: Struct., Funct., Genet.* 50, 1–4.
- (69) Holehouse, A. S., Garai, K., Lyle, N., Vitalis, A., and Pappu, R. V. (2015) Quantitative Assessments of the Distinct Contributions of Polypeptide Backbone Amides versus Side Chain Groups to Chain Expansion via Chemical Denaturation. *J. Am. Chem. Soc.* 137, 2984–2995.
- (70) Wu, Y., Kondrashkina, E., Kayatekin, C., Matthews, C. R., and Bilsel, O. (2008) Microsecond acquisition of heterogeneous structure in the folding of a TIM barrel protein. *Proc. Natl. Acad. Sci. U. S. A.* 105, 13367–13372.
- (71) Mizukami, T., Xu, M., Cheng, H., Roder, H., and Maki, K. (2013) Nonuniform chain collapse during early stages of staphylococcal nuclease folding detected by fluorescence resonance energy transfer and ultrarapid mixing methods. *Protein Sci.* 22, 1336–1348.
- (72) Rothwarf, D. M., and Scheraga, H. A. (1996) Role of non-native aromatic and hydrophobic interactions in the folding of hen egg white lysozyme. *Biochemistry* 35, 13797–13807.
- (73) Feng, H., Takei, J., Lipsitz, R., Tjandra, N., and Bai, Y. (2003) Specific Non-Native Hydrophobic Interactions in a Hidden Folding Intermediate: Implications for Protein Folding. *Biochemistry* 42, 12461–12465.
- (74) Dasgupta, A., Udgaonkar, J. B., and Das, P. (2014) Multistage Unfolding of an SH3 Domain: An Initial Urea-Filled Dry Molten Globule Precedes a Wet Molten Globule with Non-Native Structure. *J. Phys. Chem. B* 118, 6380–6392.
- (75) Chen, T., and Chan, H. S. (2015) Native Contact Density and Nonnative Hydrophobic Effects in the Folding of Bacterial Immunity Proteins. *PLoS Comput. Biol.* 11, e1004260–e1004260.
- (76) Zarrine-Afsar, A., Wallin, S., Neculai, A. M., Neudecker, P., Howell, P. L., Davidson, A. R., and Chan, H. S. (2008) Theoretical and experimental demonstration of the importance of specific nonnative interactions in protein folding. *Proc. Natl. Acad. Sci. U. S. A.* 105, 9999–10004.
- (77) Faisca, P. F. N., Nunes, A., Travasso, R. D. M., and Shakhnovich, E. I. (2010) Non-native interactions play an effective role in protein folding dynamics. *Protein Sci.* 19, 2196–2209.
- (78) Sakurai, K., Fujioka, S., Konuma, T., Yagi, M., and Goto, Y. (2011) A Circumventing Role for the Non-Native Intermediate in the Folding of  $\beta$ -Lactoglobulin. *Biochemistry* 50, 6498–6507.
- (79) Shental-Bechor, D., Smith, M. T. J., MacKenzie, D., Broom, A., Marcovitz, A., Ghashut, F., Go, C., Bralha, F., Meiering, E. M., and Levy, Y. (2012) Nonnative interactions regulate folding and switching of myristoylated protein. *Proc. Natl. Acad. Sci. U. S. A.* 109, 17839–17844.
- (80) Sosnick, T. R., Shtilerman, M. D., Mayne, L., and Englander, S. W. (1997) Ultrafast signals in protein folding and the polypeptide contracted state. *Proc. Natl. Acad. Sci. U. S. A.* 94, 8545–8550.
- (81) Qi, P. X., Sosnick, T. R., and Englander, S. W. (1998) The burst phase in ribonuclease A folding and solvent dependence of the unfolded state. *Nat. Struct. Biol.* 5, 882–884.
- (82) Chan, H. S., and Dill, K. a. (1990) Origins of structure in globular proteins. *Proc. Natl. Acad. Sci. U. S. A.* 87, 6388–6392.
- (83) Epan, R. M., and Scheraga, H. A. (1968) The influence of long-range interactions on the structure of myoglobin. *Biochemistry* 7, 2864–2872.
- (84) Baldwin, R. L. (1995)  $\alpha$ -Helix formation by peptides of defined sequence. *Biophys. Chem.* 55, 127–135.
- (85) Aghera, N., Earanna, N., and Udgaonkar, J. B. (2011) Equilibrium Unfolding Studies of Monellin: The Double-Chain Variant Appears To Be More Stable Than the Single-Chain Variant. *Biochemistry* 50, 2434–2444.
- (86) Malhotra, P., and Udgaonkar, J. B. (2014) High-Energy Intermediates in Protein Unfolding Characterized by Thiol Labeling under Nativelike Conditions. *Biochemistry* 53, 3608–3620.