THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL

Letter

Resolving Site-Specific Heterogeneity of the Unfolded State under Folding Conditions

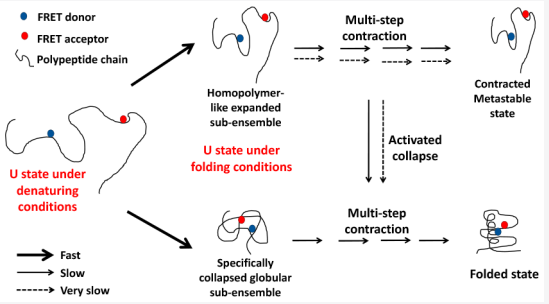
Sandhya Bhatia, Guruswamy Krishnamoorthy, and Jayant B. Udgaonkar*

Cite This: *J. Phys. Chem. Lett.* 2021, 12, 3295−3302

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Understanding the properties of the unfolded state under folding conditions is of fundamental importance for gaining mechanistic insight into folding as well as misfolding reactions. Toward achieving this objective, the folding reaction of a small protein, monellin, has been resolved structurally and temporally, with the use of the multisite time-resolved FRET methodology. The present study establishes that the initial polypeptide chain collapse is not only heterogeneous but also structurally asymmetric and nonuniform. The population-averaged size for the segments spanning parts of the β-sheet decreases much more than that for the α-helix. Multisite measurements enabled specific and nonspecific components of the initial chain collapse to be discerned. The expanded and compact intermediate subensembles have the properties of a nonspecifically collapsed (hence, random-coil-like) and specifically collapsed (hence, globular) polymer, respectively. During subsequent folding, both the subensembles underwent contraction to varying extents at the four monitored segments, which was close to gradual in nature. The expanded intermediate subensemble exhibited an additional very slow contraction, suggestive of the presence of non-native interactions that result in a higher effective viscosity slowing down intrachain motions under folding conditions.



To gain mechanistic insight into protein folding reactions, it is crucial to have an understanding of the properties of the unfolded (U) state under folding conditions, from which structure formation commences.[1,2] The conformational properties of the U state are also relevant in various other contexts such as liquid−liquid phase separation[3] and protein aggregation[4,5] and need to be understood in the context of polymer physics theory.[6,7] The U state populated under unfolding conditions has been characterized as a random-coil-like chain[8,9] with some residual structure in the case of a few proteins[10,11] and displays the properties of a polymer chain in a good solvent.[2,7−9] Intrinsically disordered proteins under denaturing conditions have also been described by a self-avoiding random walk model, as expected for a random-coil-like polymer in which chain−solvent interactions dominate over chain−chain interactions, thereby giving rise to an expanded form.[8,12,13]

Little is known about what happens to the conformational properties of the unfolded polypeptide chain upon transfer from denaturing to renaturing solvent conditions. Does it remain expanded or undergo a collapse? Does it have specific, native (N)-like interactions or nonspecific, non-native interactions? Is it globular or random-coil-like in its shape? These fundamental questions have remained a matter of active research and debate over the past few decades,[1,2,14−16] partly because the U state in folding conditions is too short-lived to be characterized by high-resolution structural probes. If chain−chain interactions dominate over the chain−solvent interactions under folding conditions, then the polypeptide chain would undergo a solvent-induced collapse, as predicted for a polymer in a poor solvent.[2,6,7,16]

Many experimental studies have probed the properties of the initially collapsed intermediate ensemble to verify the predictions of polymer theory[6,7] and to determine whether the solvent behaves as a good or poor solvent under folding conditions.[2,14,15] The results have varied considerably across proteins[2,14] and have depended on the methodologies used for monitoring the changes in protein dimensions.[17,18] Both single-molecule[8,17,19] and ensemble[2,14,20,21] FRET measurements have suggested that the U state collapses under folding conditions. However, small-angle X-ray scattering (SAXS) measurements, which measure the average radius of gyration, suggest that the U state under folding conditions remains expanded.[17,18,22] The apparent discrepancy between SAXS and smFRET-based measurements,[23] which has complicated precise determination of the extent of collapse in some cases,[17,18] may arise because of heterogeneity in the collapsed intermediate ensemble.[23]

Received: January 11, 2021
Accepted: March 23, 2021
Published: March 25, 2021

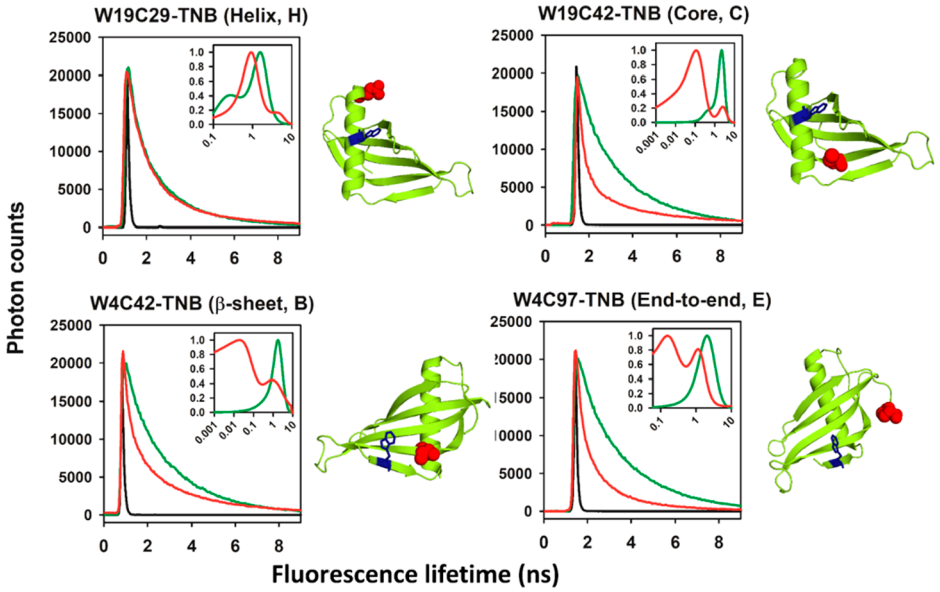


Figure 1. Multisite trFRET is a suitable probe for monitoring the folding reaction of MNEI. Time-resolved fluorescence decay traces for the TNB-labeled variants were obtained at pH 8 and 25 °C. The different panels correspond to data obtained for different labeled variants, as indicated on the top of each panel, and represent different intramolecular FRET pairs, The positions of the donor (shown as blue sticks) and acceptor (shown as red spheres) fluorophores of each FRET pair are shown in the structure on the right of each panel. The structures (for pdb ID 1IV9) were drawn using Pymol. The black, red, and green curves in each panel represent the instrument response function (IRF), the fluorescence decay trace of the native protein (in 0.4 M GdnHCl), and the fluorescence decay trace of unfolded protein (in 4 M GdnHCl), respectively. The inset in each panel shows the MEM analysis-derived fluorescence lifetime distributions for the unfolded and native proteins. The *y*-axis (amplitude) in each inset represents the relative populations for the MEM-analysis derived fluorescence lifetime histograms as a function of logarithmically spaced fluorescence lifetimes (*x*-axis). The amplitudes in each inset have been normalized to the peak amplitude for each distribution.

Some recent studies[15,24] suggest that evolution would have selected for an expanded U state even under folding conditions, because preferential chain−chain interactions that would stabilize a compact U state could also be expected to facilitate aggregation. Conversely, the formation of a collapsed (compact) U state under folding conditions[1,2,14,25] would facilitate further folding by establishing various N-like interactions and reducing the conformational entropy for folding. Experimental and computational data supporting both points of view exist.[2,14,15] Some proteins appear to undergo a very fast specific collapse leading to the formation of an N-like productive globular intermediate on the millisecond time scale, whose interactions appear to facilitate further folding.[26−30] For many other proteins, however, the collapsed form appears to be devoid of substantial specific N-like interactions.[21,31−33]

The multisite time-resolved FRET (trFRET) methodology can resolve this heterogeneity and provide comprehensive structural insight into the properties of the collapsed intermediate ensemble. trFRET provides great temporal resolution for monitoring transiently populated species,[34] while multisite measurements provide a characterization of different structural segments of the protein.[35−37] This method has been used successfully in the past to characterize the structural heterogeneity present in both equilibrium[34,36] and kinetic measurements[30,38,39] of (un)folding reactions.

Recently, the trFRET method was used for characterizing the heterogeneity in the collapse and folding of monellin (MNEI), a well-characterized model system for the study of protein folding.[37,40] Earlier multisite steady-state FRET (ssFRET) measurements had indicated that the initially collapsed intermediate, formed within the first 37 μs during folding, was stabilized by both specific and nonspecific interactions,[41,42] although it could not be ascertained if they were present in the same or different subensembles of molecules. trFRET measurements, using a single FRET pair monitoring collapse at the core (C),[40] then established that the initially collapsed intermediate ensemble is a heterogeneous mixture of expanded and compact subensembles (both distinct from the U and N states). As folding progresses, the expanded subensemble transforms into the compact subensemble in a barrier-limited cooperative manner via parallel pathways.[40] A recent multisite trFRET study[37] probing structural changes across three more segments spanning the α-helix (H), the β-sheet (B), and the end-to-end distance (E) identified multiple structurally distinct subensembles in the initially collapsed intermediate ensemble. Each subensemble was shown to utilize a distinct pathway for subsequent folding, and the sequence of structure forming events on each pathway was determined.[37]

The analysis of the data obtained in the previous study[37] was restricted to analyzing how the different initial subensembles of molecules transformed into progressively more folded subensembles in barrier-limited steps on multiple pathways, with time of folding. In that study, it had been noted that structural changes also occurred within each initial subensemble during the time it was populated, before it transited into a more structured subensemble, but these changes had not been analyzed. The present study follows up on the previous ones[37,40] by analyzing the same data obtained earlier[37] to determine the nature and extent of the structural changes occurring within the subensembles populated in the initial collapsed intermediate ensemble, with time of folding. Both the average dimensions of the collapsed intermediate, as well as the properties of the expanded and compact subensembles, have been investigated as a function of folding time and chain length. Specific and nonspecific components of the polypeptide chain collapse have been discerned, and both globular and

https://doi.org/10.1021/acs.jpclett.1c00098
J. Phys. Chem. Lett. 2021, 12, 3295−3302

random-coil-like subensembles have been shown to coexist within the same collapsed intermediate ensemble highlighting the inherent heterogeneity in the initial collapse transition.

Figure 1 shows that trFRET measurements of four intramolecular distances separating four different FRET pairs can distinguish between the N and the U states. For all the FRET pairs, the discrete analysis-derived mean fluorescence lifetime ($\tau_m = \sum \tau_i \alpha_i, \sum \alpha_i = 1$, i = 2 to 4; obtained from a multiexponential fit to the fluorescence decay traces) of the N state was smaller than that of the U state, as evident in its faster fluorescence decay, and also in the position of the peak of its maximum entropy method (MEM) analysis-derived fluorescence lifetime distribution (insets, Figure 1) (see Materials and Methods, SI). For all the FRET pairs, the fluorescence lifetime distribution for the U state was unimodal and had a peak at a fluorescence lifetime of ∼1−2 ns. The fluorescence lifetime distributions seen for the N state were found to vary across the different TNB-labeled variants corresponding to the different FRET pairs. For W19C42-TNB, W4C97-TNB, and W4C42-TNB, reporting on the core (C), end-to-end (E), and β-sheet (B) distances, respectively, the N state displayed a clear bimodal fluorescence lifetime distribution, with the major component at lower fluorescence lifetimes. For W19C29-TNB, reporting on the helix (H) segment, only a unimodal population distribution of fluorescence lifetimes was observed for both the U and N states.

The folding of MNEI was initiated by diluting the GdnHCl concentration from 4 to 0.4 M, using a stopped-flow mixer. The observed shift in the distribution to lower fluorescence lifetimes upon folding was indicative of an increase in FRET efficiency. This was confirmed by analyzing the data obtained for both the unlabeled and the TNB-labeled variants (Figures S1−S4) and determining the FRET efficiency (see Materials and Methods, SI) at each time of folding. It should be noted that the fluorescence lifetime distributions for the unlabeled protein variants were largely unimodal. Although small minor peaks at lower lifetimes were observed (Figures S1−S4) in some cases, their relative contribution to the overall population was <15% at all times of folding (Figure S5). From the FRET efficiency data, the distance separating the donor and acceptor could be determined by the use of the Forster equation.[36,43] It is important to note that mutations and labeling had no significant effect on stability, secondary structure and folding kinetics of MNEI.[36,37]

At the first observable time of folding (i.e., 100 ms), all four average intramolecular distances (<R_{DA}>), except for the end-to-end distance (in W4−C97TNB), became significantly shorter than in the U state (Figure 2a). This burst phase (<100 ms) decrease in the average intramolecular distances is referred to as the collapse transition, which is known to precede structure formation in the case of monellin.[37,40−42] The fractional change in average size ($f_{<R_{DA}>}$) at 100 ms of folding, relative to the total change in size from the U to the N state ($f_{<R_{DA}>} = \frac{<R_{DA}>^U - <R_{DA}>^t}{<R_{DA}>^U - <R_{DA}>^N}$; t = 100 ms), varied significantly (4% to 51%) between the four monitored segments, with an average value of 34% ± 10%. Previous multisite ssFRET based folding studies in the microsecond regime[42] had revealed that the initial collapse occurred as early as within 37 μs of folding, and that the size of the protein then did not change, across all the monitored segments, over the next 1 ms.[42] At 1 ms of folding, the size had undergone an average

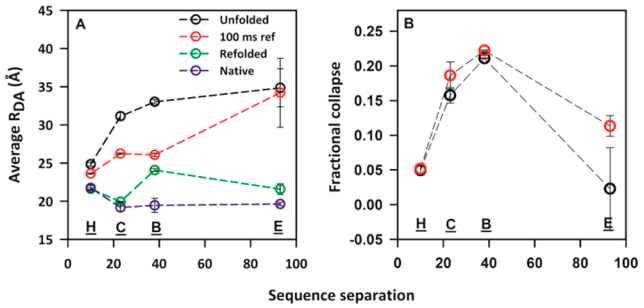


Figure 2. Asymmetric properties of the collapsed intermediate ensemble. (a) Dependence of the average donor−acceptor pair distance (<R_{DA}>), determined at different times of the folding reaction on sequence separation between the FRET donor and acceptor. The average R_{DA} was derived from mean fluorescence lifetime values obtained for both the labeled and unlabeled variants for all the FRET pairs. The distances were determined using the FRET efficiency values and the Forster distance values in the Forster equation.[36] In panel A, H, C, B, and E refer to the segments spanning the helix (W19-C29TNB), core (W19-C42TNB), β-sheet (W4-C42TNB), and end-to-end separation (W4-C97TNB). (b) Comparison of the fractional change in size with respect to the U state ($f = \frac{<R_{DA}>^U - <R_{DA}>^t}{<R_{DA}>^U}$) at 100 ms (black) and 200 ms (red) of folding for all the four intramolecular distances. The error bars represent the standard errors of measurements from two independent double kinetics experiments.

fractional change of 35% ± 9%.[40,41] It was also shown that the size decreases further only by 5−10% across various segments, and that structure formation occurs to the same extent, from 1 to 100 ms of folding.[40,41] Hence, the collapsed intermediate ensemble populated at 100 ms of folding ($f_{<R_{DA}>}$ = 34% ± 10%), studied here, resembles in its dimensions the initially collapsed intermediate ensemble populated at 37 μs of folding ($f_{<R_{DA}>}$ = 35% ± 9%). No major changes in protein dimensions seem to occur from 37 μs to 100 ms of folding.[37,40−42]

The extent of reduction in dimensions relative to the corresponding U state ($f = \frac{<R_{DA}>^U - <R_{DA}>^t}{<R_{DA}>^U}$) was different across different intramolecular distances (Figure 2b). The reduction in average size seen by trFRET in the collapsed intermediate, varied from ∼5% for the segment H spanning the α-helix, to 16% for segment C spanning the core, and 22% for segment B spanning the β-sheet. This suggested that the initial chain collapse reaction in monellin is nonuniform, as was suggested earlier using steady-state fluorescence measurements for some of these FRET pairs.[41] A computational study had also suggested that the collapse reaction of monellin is asymmetric in nature.[44] More interestingly, it was observed that the extent of the collapse was higher for segments B and C spanning the β-sheet (in W4-C42TNB and W19-C42TNB) as compared with segment H spanning the helix (in W19-C29TNB) (Figure 2b). This is also in concordance with a recent study suggesting more collapsibility in β-sheet rich proteins than α-helix containing ones,[14] and also provides experimental support to the prediction that the extent of collapse depends upon the topology of the corresponding folded region in a protein.[45]

Figure 3 shows the distance distributions measured for different structural segments, which were determined from the

https://doi.org/10.1021/acs.jpclett.1c00098
J. Phys. Chem. Lett. 2021, 12, 3295−3302

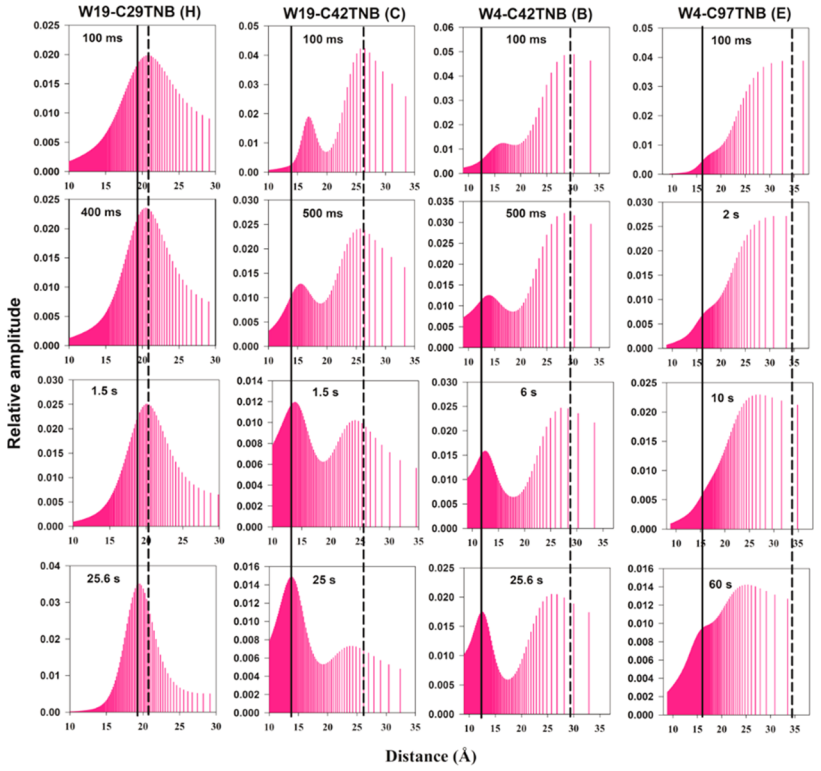


Figure 3. Evolution of distance distributions as a function of folding time. Experimentally derived distance distributions for representative time points of the folding reaction. The topmost panel corresponds to the U state, the bottom-most panel corresponds to the distance distribution of the refolded N state, and the middle panels correspond to intermediate time points, as described on each panel, during folding following a 4 to 0.4 M GdnHCl jump. The vertical solid and dashed black lines indicate the peak positions of the fluorescence lifetime distributions corresponding to the refolded N state and U state, respectively.

MEM analysis of trFRET data. The distance distributions were derived from the fluorescence lifetime distributions of the TNB-labeled variants upon accounting for the changes in the corresponding unlabeled variants (Figures S1−S4), as described earlier[40] (also see Materials and Methods section, SI). This step ensured that other biophysical effects that might affect the fluorescence lifetime distributions due to changes in the Trp environment during folding were taken care of. In the previous study probing collapse and folding only at segment C,[40] a bimodal distance distribution was observed for the burst phase intermediate populated at 100 ms of folding. Figure 3 shows that the collapsed intermediate ensemble at 100 ms of folding is similarly heterogeneous at segments B and E as well. It is important to note that such bimodality is seen only for the TNB-labeled variants probing FRET and not for the corresponding unlabeled variants probing changes in the Trp environment (Figures S2−S4). This result emphasizes the power of MEM analysis to uniquely and robustly resolve the heterogeneity from ensemble trFRET measurements.[34] The two subpopulations observed for the collapsed intermediate ensemble at segments C, B, and E were identified as U-like (R_{DA} > 20 Å) and N-like (R_{DA} < 20 Å) subensembles, based upon their similarities in MEM-peak distance with the U and N states,[37,40] respectively (Figures 3 and 4). This result is in accordance with theoretical predictions of kinetic partitioning of folding molecules immediately after initial chain collapse.[46]

A homopolymer in a θ or good solvent is expected to undergo an expansion in size with an increase in length.[2,6,7,47] Structural segments that remained U-like at 100 ms of folding were found to show a monotonic increase in spatial separation

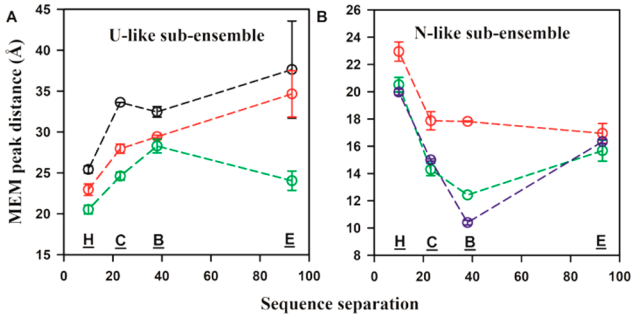


Figure 4. The U-like subensemble has properties of a homopolymer akin to the U state, while the N-like subensemble is globular like the N state. Dependence of the spatial separation (R_{DA}) on the sequence separation between the FRET donor and acceptor, measured for the U-like structural segments (panel a) and the N-like structural segments (panel b), at different times of folding. The peak distance, at different times of folding, was obtained using the most probable fluorescence lifetimes of the MEM-derived fluorescence lifetime distributions for the TNB-labeled variants and their corresponding unlabeled counterparts (Figures S1−S4).[40] Different colors correspond to data for the U state in 4 M GdnHCl (black), at 100 ms of folding (red), refolded state (green), and the N state (blue). H, C, B, and E refer to different segments, as described in Figure 2. The error bars correspond to the standard errors measured from two independent double kinetics experiments.

(intramolecular distance, R_{DA}) with increasing sequence separation (Figure 4A). In the case of the unfolded protein in 4 M GdnHCl, the increase was not strictly monotonic but close to being so, and the upward trend was similar to that seen

https://doi.org/10.1021/acs.jpclett.1c00098
J. Phys. Chem. Lett. 2021, 12, 3295−3302

for the U-like subensemble at 100 ms of folding. The observation that both the U state in 4 M GdnHCl and the U-like subensemble at 100 ms of folding show similar positive correlations between the measured R_{DA} and the length of the probed segment (Figure 4A) confirms the homopolymer-like properties of the U-like subensemble.[47] It suggests that the contraction from the U state to a U-like subensemble involves only nonspecific intrachain interactions,[20,21,24,48] and thus, the random-coil-like properties of the U state (Figures 2A and 4A) are retained in the U-like subensemble at 100 ms of folding. A determination of the scaling exponent describing the dependence of size on chain length would reveal details about the solvent quality for the observed homopolymer-like behavior.[6,7,47] Unfortunately, the R_{DA} measured using FRET cannot be converted accurately to the radius of gyration, and hence, the solvent quality cannot be interpreted.[23]

In contrast, structural segments that have collapsed to become N-like at 100 ms of folding did not show such a monotonic increase in spatial separation with increasing sequence separation (Figure 4b). Instead, the dependence of spatial separation on sequence separation for the N-like structural segments was more similar to that observed for the globular N state (Figure 4b). This confirmed the presence of at least some specific native-like interactions[29−33,49] in the N-like compact and globular subensemble. Although no correlation of spatial separation with sequence separation was observed previously in the case of the initial kinetic intermediate of barstar,[20,49] that was likely to have been the consequence of an inability, in ssFRET measurements, to resolve between expanded (such as U-like) and compact (such as N-like) structural segments in the initial kinetic intermediate ensemble.

A continuous movement in the peak position of the distance distributions toward a shorter distance was seen at all the four structural segments monitored by FRET (Figures 3 and 5). This was observed for the N-like distance distribution at segment B, for the U-like distance distribution at segment E, and both the N-like and U-like distance distributions at segment C (Figure 5). A continuous shift was also observed for the distance distribution monitoring the sequence segment spanning the helix in all protein molecules, whether in U-like or N-like subensembles, as the folding reaction progressed. Unfortunately the magnitudes of these continuous shifts in the peaks of the distance distributions were small compared with the widths of the distributions (Figures 3 and S6). It was therefore not possible to definitively conclude that these distances contract in a continuous barrier-less manner. Nevertheless, the observation of very many crossover points of all the distance distributions (Figure S6), suggests strongly that the transition occurs through a multitude of states, and is close to being continuous in nature. Future studies of the temperature dependence of the transition will establish whether the transition is truly continuous.[20]

In this context, it is important to point out that a barrier-less continuous contraction can be both exponential in its kinetics and slow. Theoretical[50] as well as experimental[20,39,40,51] studies have indicated that diffusive motions in polypeptide chains accompanying changes in size can result in nearly exponential relaxation kinetics, especially if the chain motions are slow, reflective of a rugged underlying landscape.[50,52] The slowness in chain motions could be because: (a) The energy landscape is rugged, that is, folding would involve the crossing of multiple small (<3 k_BT) free energy barriers arising from imperfect enthalpy−entropy compensations, resulting in multistep

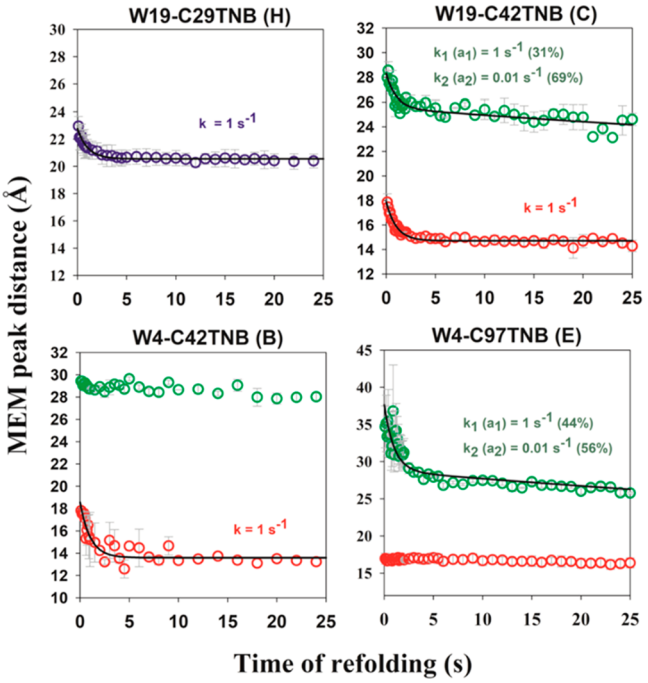


Figure 5. Continuous movement of the peaks of the fluorescence lifetime distributions. Kinetics of contraction of U-like (green empty circles) and N-like (red empty circles) distance distributions in the different subensembles, as determined for the different FRET pair variants (as mentioned on top of each panel) in 0.4 M GdnHCl at pH 8 and 25 °C. The peak distance, at different times of folding, was obtained using the most probable fluorescence lifetimes of the MEM derived fluorescence lifetime distributions for the TNB-labeled variants and their corresponding unlabeled counterparts (Figures S1−S4).[40] For the labeled variant mapping the helix (W19C29-TNB), only a unimodal peak was observed (Figure S1), and the data for the same is shown in blue color. The solid lines through the kinetics data are fit to either a single or double exponential equation. The rate constants and relative amplitudes for the observed exponential kinetic phases are reported in the individual panels. The error bars represent the standard errors of measurements from two independent double kinetics experiments.

conformational changes;[53] (b) Intrachain motions could be slaved to solvent motions[54] and could be dampened more for a collapsed intermediate ensemble, possibly because of associated steric hindrance; and (c) Further contraction involves the formation of new intramolecular contacts with high effective activation free energy.

In a previous study,[40] the unusually slow contraction of one intramolecular distance was identified as glassy-like behavior. This was based upon viscosity calculations that suggested that the contraction was slowed down by an effective viscosity that had a value 10^9 times larger than that of water (see Materials and Methods in the SI). In this study, it is shown that three other intramolecular distances show similarly slow contraction. The dynamics observed in glassy states are poorly understood in terms of the barriers that slow them down drastically. Transitions within glassy states can be continuous and barrier-less (<3 k_BT),[55,56] or they can be barrier-limited with Arrhenius or super-Arrhenius kinetics.[57−60] At present it is not known whether the glassy-like kinetics observed in this study are barrier-less or barrier-limited. This will be the focus of future studies (see above).

https://doi.org/10.1021/acs.jpclett.1c00098
J. Phys. Chem. Lett. 2021, 12, 3295−3302

Interestingly, the U-like subensemble underwent contraction only at segments C and E and not at the segment B. It could be that the contraction of the β-sheet segment is retarded by the presence of non-native interactions in the U-like subensemble[48,61] and that contraction of this segment becomes possible only when these non-native interactions get broken in a barrier-limited step, corresponding to the U-like to N-like conversion during folding.[37,40] Figure 5 also shows that the N-like subensemble contracted with the time of folding only at segments B and C; no contraction was observed at segment E. The contraction at segment E could have been prevented by native as well as non-native interactions that form in other segments of the structure.[28,29] The current understanding of the origin of the observed nonuniform and close to gradual nature of the contraction of different structural segments is far from satisfactory and will be the focus of future studies.

The U-like subensemble underwent contraction at all structural segments more slowly (with an additional 0.01 s^{-1} rate constant) than the N-like subensemble (1 s^{-1}) (Figure 5). It appears that the two subensembles differ in the extent of non-native interactions present. Fewer (if any) such interactions present in the specifically collapsed N-like subensemble are likely to retard its further contraction. On the other hand, more non-native interactions present in the expanded and nonspecifically collapsed random-coil-like U-like subensemble, would retard its further contraction. In other words, non-native interactions would increase the effective viscosity experienced by the polymer chain[51,62,63] during its diffusion over a more rugged free energy landscape. It should be noted that the rate constants reported in Figure 5 are only rough estimates of the true rate constants of contraction, as peak shift is not proportional to population. These estimates serve the purpose of establishing the similarities and differences in the time scales of contraction of different distances in the U-like and N-like subensembles.

The U-like subensemble seems to correspond to a contracted and stabilized version of the U state ensemble that underwent solvent-induced chain readjustment upon transfer to folding conditions.[31−33] The stabilization with respect to the U-state might be the reason for increased ruggedness and effective viscosity leading to slow barrier-limited[37] as well as close to barrier-less changes during folding.[31,40] However, the N-like subensemble seems to correspond to a globular and productive on-pathway folding intermediate[1,26,27] that contains various native-like interactions that facilitate further folding to the N-state.[40]

Earlier multisite trFRET studies, carried out under equilibrium conditions,[36] had shown that segment H displayed a unimodal population distribution of intramolecular distances whose peak shifted continuously with a decrease in GdnHCl concentration. Segments B, C, and E showed both U-like and N-like population distributions, with the former converting into the latter in an all-or-none manner, while both distributions contracted in a close to continuous manner with a decrease in GdnHCl concentration. The equilibrium results had suggested that the U-like and N-like subensembles, present at any time during folding, consisted of molecules that differed in their degree of solvent exposure. The addition of GdnHCl would preferentially stabilize the more expanded molecules in each subensemble. The observation that each subensemble contracted in a close to continuous manner upon a reduction in GdnHCl concentration suggested that the free energy barriers separating compact and expanded molecules in a partly (un)folded subensemble were small. Similar results had been seen in earlier equilibrium studies of the unfolding of barstar[34] and the SH3 domain of PI3 kinase.[64] This inference from equilibrium unfolding studies is supported by multisite trFRET-monitored kinetic studies of the unfolding of MNEI.[39] Those studies showed that segments B and E (H and C were not monitored) expanded continuously with time of unfolding,[39] suggesting that unfolding proceeded through a slow and diffusive swelling of molecules in N-like and U-like subensembles and that the former converted into the latter through the barrier-limited breakage of a few noncovalent interactions. The slow and close to continuous changes during folding, observed in four different segments in this study, have been shown to be describable by a modified Rouse-like model of polymer physics,[40] as are the similar changes observed previously in kinetic unfolding studies.[39]

The present study has revealed and quantified segment-specific, structural heterogeneity in the initially collapsed form populated during the folding of MNEI. The segment spanning the α-helix undergoes a close to continuous slow reduction in size, homogeneously across all molecules, during folding. However, for other parts of the protein such as the core, β-sheet, and end-to-end separation, the collapsed intermediate ensemble is heterogeneous. It consists of a random-coil-like nonspecifically collapsed U-like subensemble and a specifically collapsed N-like subensemble. The collapsed intermediate ensemble undergoes a nonuniform glassy-like slow contraction at different structural segments at rates that are significantly different for the U-like and N-like subensembles.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.1c00098.

Materials and Methods and supplementary figures (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jayant B. Udgaonkar − *National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru 560 065, India; Indian Institute of Science Education and Research, Pune 411 008, India;* orcid.org/0000-0002-7005-224X; Email: jayant@ncbs.res.in, jayant@iiserpune.ac.in

Authors

Sandhya Bhatia − *National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru 560 065, India; Indian Institute of Science Education and Research, Pune 411 008, India;* orcid.org/0000-0002-0926-8287

Guruswamy Krishnamoorthy − *Department of Biotechnology, Anna University, Chennai 600 025, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.1c00098

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Deepak Dhar for discussions regarding glassy transitions. We would also like to thank members of our

laboratory, as well as M. K. Mathew and S. Gosavi, for discussion. J.B.U. is a recipient of a JC Bose National Fellowship from the Government of India. This work was funded by the Tata Institute of Fundamental Research, and by the Department of Science and Technology, Government of India.

REFERENCES

(1) Chan, H. S.; Dill, K. A. Origins of Structure in Globular Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87* (16), 6388−6392.
(2) Udgaonkar, J. B. Polypeptide Chain Collapse and Protein Folding. *Arch. Biochem. Biophys.* **2013**, *531* (1−2), 24−33.
(3) Uversky, V. N.; Finkelstein, A. V. Life in Phases: Intra-and Inter-Molecular Phase Transitions in Protein Solutions. *Biomolecules* **2019**, *9* (12), 842.
(4) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. Measuring the Rate of Intramolecular Contact Formation in Polypeptides. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (13), 7220−7225.
(5) Gerum, C.; Schlepckow, K.; Schwalbe, H. The Unfolded State of the Murine Prion Protein and Properties of Single-Point Mutants Related to Human Prion Diseases. *J. Mol. Biol.* **2010**, *401* (1), 7−12.
(6) Flory, P. J. *Principles of Polymer Chemistry*; Cornell University Press: Ithaca, NY, 1953.
(7) de Gennes, P. G. Collapse of a Polymer-Chain in Poor Solvents. *J. Phys., Lett.* **1975**, *36* (3), L55−L57.
(8) Hofmann, H.; Soranno, A.; Borgia, A.; Gast, K.; Nettels, D.; Schuler, B. Polymer Scaling Laws of Unfolded and Intrinsically Disordered Proteins Quantified with Single-Molecule Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (40), 16155−16160.
(9) Borgia, A.; Zheng, W.; Buholzer, K.; Borgia, M. B.; Schüler, A.; Hofmann, H.; Soranno, A.; Nettels, D.; Gast, K.; Grishaev, A.; et al. Consistent View of Polypeptide Chain Expansion in Chemical Denaturants from Multiple Experimental Methods. *J. Am. Chem. Soc.* **2016**, *138* (36), 11714−11726.
(10) Mok, K. H.; Kuhn, L. T.; Goez, M.; Day, I. J.; Lin, J. C.; Andersen, N. H.; Hore, P. J. A Pre-Existing Hydrophobic Collapse in the Unfolded State of an Ultrafast Folding Protein. *Nature* **2007**, *447* (7140), 106−109.
(11) Ota, C.; Ikeguchi, M.; Tanaka, A.; Hamada, D. Residual Structures in the Unfolded State of Starch-Binding Domain of Glucoamylase Revealed by near-UV Circular Dichroism and Protein Engineering Techniques. *Biochim. Biophys. Acta, Proteins Proteomics* **2016**, *1864* (10), 1464−1472.
(12) Grupi, A.; Haas, E. Segmental Conformational Disorder and Dynamics in the Intrinsically Disordered Protein α-Synuclein and Its Chain Length Dependence. *J. Mol. Biol.* **2011**, *405* (5), 1267−1283.
(13) Zheng, W.; Borgia, A.; Buholzer, K.; Grishaev, A.; Schuler, B.; Best, R. B. Probing the Action of Chemical Denaturant on an Intrinsically Disordered Protein by Simulation and Experiment. *J. Am. Chem. Soc.* **2016**, *138* (36), 11702−11713.
(14) Thirumalai, D.; Samanta, H. S.; Maity, H.; Reddy, G. Universal Nature of Collapsibility in the Context of Protein Folding and Evolution. *Trends Biochem. Sci.* **2019**, *44*, 675−687.
(15) Clark, P. L.; Plaxco, K. W.; Sosnick, T. R. Water as a Good Solvent for Unfolded Proteins: Folding and Collapse Are Fundamentally Different. *J. Mol. Biol.* **2020**, *432* (9), 2882−2889.
(16) Graziano, G. Is Water a Good Solvent for the Denatured State of Globular Proteins? *Chem. Phys. Lett.* **2020**, *759*, 137949.
(17) Yoo, T. Y.; Meisburger, S. P.; Hinshaw, J.; Pollack, L.; Haran, G.; Sosnick, T. R.; Plaxco, K. Small-Angle X-Ray Scattering and Single-Molecule FRET Spectroscopy Produce Highly Divergent Views of the Low-Denaturant Unfolded State. *J. Mol. Biol.* **2012**, *418* (3−4), 226−236.
(18) Watkins, H. M.; Simon, A. J.; Sosnick, T. R.; Lipman, E. A.; Hjelm, R. P.; Plaxco, K. W. Random Coil Negative Control Reproduces the Discrepancy between Scattering and FRET Measurements of Denatured Protein Dimensions. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (21), 6631−6636.
(19) Hoffmann, A.; Kane, A.; Nettels, D.; Hertzog, D. E.; Baumgärtel, P.; Lengefeld, J.; Reichardt, G.; Horsley, D. A.; Seckler, R.; Bakajin, O.; Schuler.; et al. Mapping Protein Collapse with Single-Molecule Flourescence and Kinetic Synchrotron Radiation Circular Dichroism Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (1), 105−110.
(20) Sinha, K. K.; Udgaonkar, J. B. Barrierless Evolution of Structure during the Submillisecond Refolding Reaction of a Small Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (23), 7998−8003.
(21) Dasgupta, A.; Udgaonkar, J. B. Evidence for Initial Non-Specific Polypeptide Chain Collapse During the Refolding of the SH3 Domain of PI3 Kinase. *J. Mol. Biol.* **2010**, *403* (3), 430−445.
(22) Riback, J. A.; Bowman, M. A.; Zmyslowski, A. M.; Knoverek, C. R.; Jumper, J. M.; Hinshaw, J. R.; Kaye, E. B.; Freed, K. F.; Clark, P. L.; Sosnick, T. R. Are Expanded in Water. *Science* **2017**, *358* (6360), 238−241.
(23) Song, J.; Gomes, G. N.; Shi, T.; Gradinaru, C. C.; Chan, H. S. Conformational Heterogeneity and FRET Data Interpretation for Dimensions of Unfolded Proteins. *Biophys. J.* **2017**, *113* (5), 1012−1024.
(24) Bowman, M. A.; Riback, J. A.; Rodriguez, A.; Guo, H.; Li, J.; Sosnick, T. R.; Clark, P. L. Properties of Protein Unfolded States Suggest Broad Selection for Expanded Conformational Ensembles. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (38), 23356−23364.
(25) Kirmizialtin, S.; Pitici, F.; Cardenas, A. E.; Elber, R.; Thirumalai, D. Dramatic Shape Changes Occur as Cytochrome c Folds. *J. Phys. Chem. B* **2020**, *124* (38), 8240−8248.
(26) Uzawa, T.; Akiyama, S.; Kimura, T.; Takahashi, S.; Ishimori, K.; Morishima, I.; Fujisawa, T. Collapse and Search Dynamics of Apomyoglobin Folding Revealed by Submillisecond Observations of α-Helical Content and Compactness. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (5), 1171−1176.
(27) Kimura, T.; Akiyama, S.; Uzawa, T.; Ishimori, K.; Morishima, I.; Fujisawa, T.; Takahashi, S. Specifically Collapsed Intermediate in the Early Stage of the Folding of Ribonuclease A. *J. Mol. Biol.* **2005**, *350* (2), 349−362.
(28) Bartlett, A. I.; Radford, S. E. Desolvation and Development of Specific Hydrophobic Core Packing during Im7 Folding. *J. Mol. Biol.* **2010**, *396* (5), 1329−1345.
(29) Fazelinia, H.; Xu, M.; Cheng, H.; Roder, H. Ultrafast Hydrogen Exchange Reveals Specific Structural Events during the Initial Stages of Folding of Cytochrome C. *J. Am. Chem. Soc.* **2014**, *136* (2), 733−740.
(30) Peran, I.; Holehouse, A. S.; Carrico, I. S.; Pappu, R. V.; Bilsel, O.; Raleigh, D. P. Unfolded States under Folding Conditions Accommodate Sequence-Specific Conformational Preferences with Random Coil-like Dimensions. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 12301−12310.
(31) Chan, C. K.; Hu, Y.; Takahashi, S.; Rousseau, D. L.; Eaton, W. A.; Hofrichter, J. Submillisecond Protein Folding Kinetics Studied by Ultrarapid Mixing. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (5), 1779−1784.
(32) Qi, P. X.; Sosnick, T. R.; Englander, S. W. The Burst Phase in Ribonuclease a Folding and Solvent Dependence of the Unfolded State. *Nat. Struct. Biol.* **1998**, *5* (10), 882−884.
(33) Magg, C.; Kubelka, J.; Holtermann, G.; Haas, E.; Schmid, F. X. Specificity of the Initial Collapse in the Folding of the Cold Shock Protein. *J. Mol. Biol.* **2006**, *360* (5), 1067−1080.
(34) Lakshmikanth, G. S.; Sridevi, K.; Krishnamoorthy, G.; Udgaonkar, J. B. Structure Is Lost Incrementally during the Unfolding of Barstar. *Nat. Struct. Biol.* **2001**, *8* (9), 799−804.
(35) Lillo, M. P.; Szpikowska, B. K.; Mas, M. T.; Sutin, J. D.; Beechem, J. M. Real-Time Measurement of Multiple Intramolecular Distances during Protein Folding Reactions: A Multisite Stopped-Flow Fluorescence Energy- Transfer Study of Yeast Phosphoglycerate Kinase. *Biochemistry* **1997**, *36* (37), 11273−11281.
(36) Bhatia, S.; Krishnamoorthy, G.; Udgaonkar, J. B. Site-Specific Time-Resolved FRET Reveals Local Variations in the Unfolding

https://doi.org/10.1021/acs.jpclett.1c00098
J. Phys. Chem. Lett. 2021, 12, 3295−3302

Mechanism in an Apparently Two-State Protein Unfolding Transition. *Phys. Chem. Chem. Phys.* **2018**, *20* (5), 3216−3232.
(37) Bhatia, S.; Krishnamoorthy, G.; Udgaonkar, J. B. Mapping Distinct Sequences of Structure Formation Differentiating Multiple Folding Pathways of a Small Protein. *J. Am. Chem. Soc.* **2021**, *143*, 1447.
(38) Ratner, V.; Amir, D.; Kahana, E.; Haas, E. Fast Collapse but Slow Formation of Secondary Structure Elements in the Refolding Transition of E. Coli Adenylate Kinase. *J. Mol. Biol.* **2005**, *352* (3), 683−699.
(39) Jha, S. K.; Dhar, D.; Krishnamoorthy, G.; Udgaonkar, J. B. Continuous Dissolution of Structure during the Unfolding of a Small Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (27), 11113−11118.
(40) Bhatia, S.; Krishnamoorthy, G.; Dhar, D.; Udgaonkar, J. B. Observation of Continuous Contraction and a Metastable Misfolded State during the Collapse and Folding of a Small Protein. *J. Mol. Biol.* **2019**, *431* (19), 3814−3826.
(41) Goluguri, R. R.; Udgaonkar, J. B. Rise of the Helix from a Collapsed Globule during the Folding of Monellin. *Biochemistry* **2015**, *54* (34), 5356−5365.
(42) Goluguri, R. R.; Udgaonkar, J. B. Microsecond Rearrangements of Hydrophobic Clusters in an Initially Collapsed Globule Prime Structure Formation during the Folding of a Small Protein. *J. Mol. Biol.* **2016**, *428* (15), 3102−3117.
(43) Lakowicz, J. R.; Kuśba, J.; Szmacinski, H.; Gryczynski, I.; Eis, P. S.; Wiczk, W.; Johnson, M. L. Resolution of End-to-end Diffusion Coefficients and Distance Distributions of Flexible Molecules Using Fluorescent Donor-acceptor and Donor-quencher Pairs. *Biopolymers* **1991**, *31* (12), 1363−1378.
(44) Maity, H.; Reddy, G. Thermodynamics and Kinetics of Single-Chain Monellin Folding with Structural Insights into Specific Collapse in the Denatured State Ensemble. *J. Mol. Biol.* **2017**, *430*, 465−478.
(45) Samanta, H. S.; Zhuravlev, P. I.; Hinczewski, M.; Hori, N.; Chakrabarti, S.; Thirumalai, D. Protein Collapse Is Encoded in the Folded State Architecture. *Soft Matter* **2017**, *13* (19), 3622−3638.
(46) Guo, Z.; Thirumalai, D. Kinetics of Protein Folding: Nucleation Mechanism, Time Scales, and Pathways. *Biopolymers* **1995**, *36* (1), 83−102.
(47) Halperin, A.; Goldbart, P. M. Early Stages of Homopolymer Collapse. *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.* **2000**, *61* (1), 565−573.
(48) Meng, W.; Lyle, N.; Luan, B.; Raleigh, D. P.; Pappu, R. V. Experiments and Simulations Show How Long-Range Contacts Can Form in Expanded Unfolded Proteins with Negligible Secondary Structure. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (6), 2123−2128.
(49) Saxena, A. M.; Udgaonkar, J. B.; Krishnamoorthy, G. Characterization of Intra-Molecular Distances and Site-Specific Dynamics in Chemically Unfolded Barstar: Evidence for Denaturant-Dependent Non-Random Structure. *J. Mol. Biol.* **2006**, *359* (1), 174−189.
(50) Hagen, S. J. Exponential Decay Kinetics in "Downhill" Protein Folding. *Proteins: Struct., Funct., Genet.* **2003**, *50* (1), 1−4.
(51) Chung, H. S.; Piana-Agostinetti, S.; Shaw, D. E.; Eaton, W. A. Structural Origin of Slow Diffusion in Protein Folding. *Science (Washington, DC, U. S.)* **2015**, *349* (6255), 1504−1510.
(52) Chavez, L. L.; Onuchic, J. N.; Clementi, C. Quantifying the Roughness on the Free Energy Landscape: Entropic Bottlenecks and Protein Folding Rates. *J. Am. Chem. Soc.* **2004**, *126* (27), 8426−8432.
(53) Zwanzig, R. Diffusion in a Rough Potential. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85* (7), 2029−2030.
(54) Frauenfelder, H.; Fenimore, P. W.; Chen, G.; McMahon, B. H. Protein Folding Is Slaved to Solvent Motions. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (42), 15469−15472.
(55) Kussell, E.; Shakhnovich, E. I. Glassy Dynamics of Side-Chain Ordering in a Simple Model of Protein Folding. *Phys. Rev. Lett.* **2002**, *89* (16), 168101.
(56) Welch, R. C.; Smith, J. R.; Potuzak, M.; Guo, X.; Bowden, B. F.; Kiczenski, T. J.; Allan, D. C.; King, E. A.; Ellison, A. J.; Mauro, J. C. Dynamics of Glass Relaxation at Room Temperature. *Phys. Rev. Lett.* **2013**, *110* (26), 265901.
(57) Stillinger, F. H.; Debenedetti, P. G. Glass Transition Thermodynamics and Kinetics. *Annu. Rev. Condens. Matter Phys.* **2013**, *4* (1), 263−285.
(58) Rodríguez-Tinoco, C.; Ràfols-Ribé, J.; González-Silveira, M.; Rodríguez-Viejo, J. Relaxation Dynamics of Glasses along a Wide Stability and Temperature Range. *Sci. Rep.* **2016**, *6*, 35607.
(59) Singh, A.; Singh, Y. Super-Arrhenius Behavior of Molecular Glass Formers. *Phys. Rev. E* **2019**, *99* (3), 030101.
(60) Ginzburg, V. V. A Simple Mean-Field Model of Glassy Dynamics and Glass Transition. *Soft Matter* **2020**, *16* (3), 810−825.
(61) Takahashi, S.; Yoshida, A.; Oikawa, H. Hypothesis: Structural Heterogeneity of the Unfolded Proteins Originating from the Coupling of the Local Clusters and the Long-Range Distance Distribution. *Biophys. Rev.* **2018**, *10* (2), 363−373.
(62) Waldauer, S. A.; Bakajin, O.; Lapidus, L. J. Extremely Slow Intramolecular Diffusion in Unfolded Protein L. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (31), 13713−13717.
(63) Balacescu, L.; Schrader, T. E.; Radulescu, A.; Zolnierczuk, P.; Holderer, O.; Pasini, S.; Fitter, J.; Stadler, A. M. Transition between Protein-like and Polymer-like Dynamic Behavior: Internal Friction in Unfolded Apomyoglobin Depends on Denaturing Conditions. *Sci. Rep.* **2020**, *10* (1), 1570.
(64) Kishore, M.; Krishnamoorthy, G.; Udgaonkar, J. B. Critical Evaluation of the Two-State Model Describing the Equilibrium Unfolding of the PI3K SH3 Domain by Time-Resolved Fluorescence Resonance Energy Transfer. *Biochemistry* **2013**, *52* (52), 9482−9496.

https://doi.org/10.1021/acs.jpclett.1c00098
J. Phys. Chem. Lett. 2021, 12, 3295−3302