CHEMICAL REVIEWS

pubs.acs.org/CR

Review

Heterogeneity in Protein Folding and Unfolding Reactions

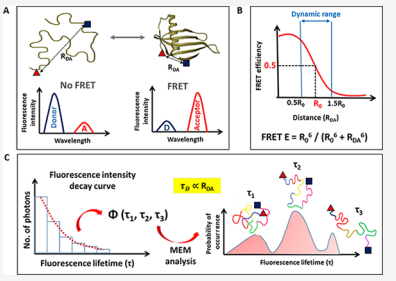
Sandhya Bhatia and Jayant B. Udgaonkar*

Cite This: https://doi.org/10.1021/acs.chemrev.1c00704

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Proteins have dynamic structures that undergo chain motions on time scales spanning from picoseconds to seconds. Resolving the resultant conformational heterogeneity is essential for gaining accurate insight into fundamental mechanistic aspects of the protein folding reaction. The use of high-resolution structural probes, sensitive to population distributions, has begun to enable the resolution of site-specific conformational heterogeneity at different stages of the folding reaction. Different states populated during protein folding, including the unfolded state, collapsed intermediate states, and even the native state, are found to possess significant conformational heterogeneity. Heterogeneity in protein folding and unfolding reactions originates from the reduced cooperativity of various kinds of physicochemical interactions between various structural elements of a protein, and between a protein and solvent. Heterogeneity may arise because of functional or evolutionary constraints. Conformational substates within the unfolded state and the collapsed intermediates that exchange at rates slower than the subsequent folding steps give rise to heterogeneity on the protein folding pathways. Multiple folding pathways are likely to represent distinct sequences of structure formation. Insight into the nature of the energy barriers separating different conformational states populated during (un)folding can also be obtained by resolving heterogeneity.



CONTENTS

Received: August 10, 2021

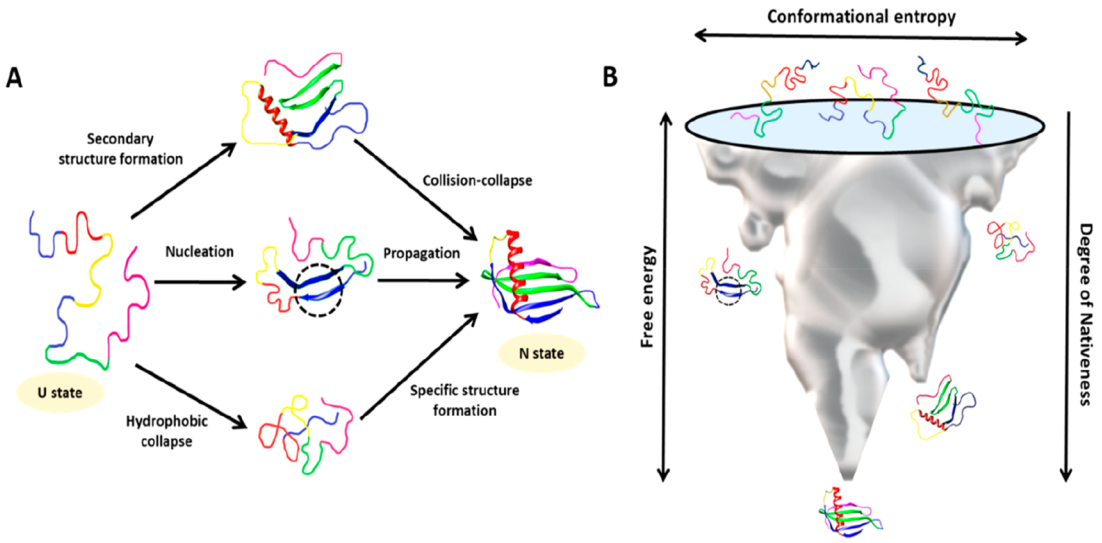


Figure 1. Models for protein folding. (A) Phenomenological models describing possible sequences of structural events associated with the transformation of the unfolded state to the specifically structured native state. The framework model (topmost pathway) posits that secondary structure formation precedes tertiary structure formation. The hydrophobic collapse model (bottom-most pathway) suggests that the initial nonspecific collapse of hydrophobic groups in the polypeptide chain drives folding. The nucleation-condensation model (central pathway) ascribes the formation of an activated folding nucleus before the transition state to be the initial step in folding. (B) The folding funnel model is described by the energy landscape theory. The width of the funnel represents the conformational entropy, the depth or height of the funnel corresponds to the free energy of the system (inclusive of interactions and solvation but excluding the contributions from the conformational entropy), and the angular coordinates of the funnel symbolize different reaction coordinates. Unfolded conformations with high entropy are present at the top of the funnel, compact partially structured intermediates are present at local minima, and the N state is present at the bottom of the funnel corresponding to the global free energy minimum. The funnel model accommodates the results of many experimental studies of protein folding.

1. INTRODUCTION

1.1. Perspectives and Overview

Proteins are the workhorse molecules that play crucial structural and functional roles in the cellular and organismal processes that define life. They are synthesized inside the cell as unstructured chains (polypeptide chains) of amino acid residues connected covalently via peptide bonds. The polypeptide chain can undergo a self-assembly process, from an unstructured unfolded (U) state to the uniquely structured native (N) state (Figure 1).[1−3] The folding process occurs within microseconds to minutes for most proteins.[4] The folded state is stabilized by a variety of intramolecular interactions that are weak and noncovalent in nature. Unlike a simple chemical reaction that involves the formation or breakage of covalent bonds, the protein folding reaction is a "disorder to order" transition, in which a large reduction in the conformational entropy of the polymer chain is compensated for by the formation of multiple noncovalent interactions,[3] as well as by favorable entropy changes in the solvent. The coupling between many different interactions defines the free energy landscape for any given protein[5] and, importantly, the degree of cooperativity in the folding process.[6,7]

1.2. The Protein Folding Problem and Its Relevance

Research on the protein folding problem focuses on understanding the rules by which various physicochemical interactions are established during the folding process.[4,8−11] It is known that under stressed cellular conditions, or due to changes in the amino acid sequence (mutations), the polypeptide chain may fail to undergo correct folding (to the N state), resulting in the formation of misfolded conformations.[12,13] These misfolded states may aggregate *in vivo*, and protein aggregation is linked to several neurodegenerative disorders.[14−17] Understanding how a protein folds will provide clues as to how to prevent misfolding and aggregation. A mechanistic understanding of the protein folding reaction will also enable the design of proteins with desired structural and unique functional traits.[18]

This review focuses on the experimental characterization of protein folding reactions. Research on the mechanism of protein folding has benefited tremendously from theoretical and computational methods, as well as from the marriage of these methods with experimental methods. This review will, however, not discuss the results of computational and theoretical methods in detail but will discuss them only in the context of providing a proper perspective on the results of experimental studies.

1.3. Phenomenological Frameworks for Protein Folding

Thinking about the results from experimental studies of protein folding, and the consequent progress in the experimental characterization of protein folding reactions, continues to be guided by several phenomenological models (Figure 1a) that describe alternative ways in which structure may develop as a protein folds.[19,20]

(1) *Framework model:* This hierarchical model posits that local secondary structural elements form early during folding, which assemble to form the tertiary structure.[21−23] Stabilization of the secondary structural units occurs when they associate upon random diffusive collision.[24,25] An underlying assumption is that coalescence of the secondary structural units can happen in many different ways. The model is a recognition of the observation that protein tertiary structures appear to be assembled from their parts.[26−28] The formation of a secondary structure early during folding was observed in many early experimental studies,[29−33] as well as in computer simulations[34,35] of folding. More recently, basic ideas of the framework model have evolved into

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

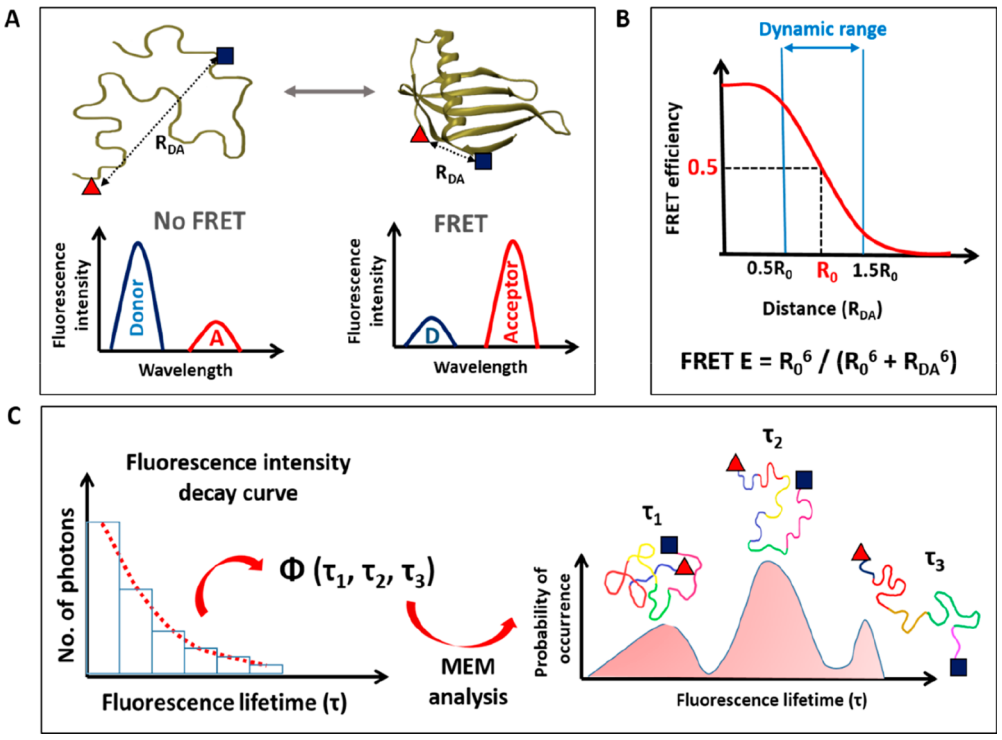


Figure 2. Resolving heterogeneity using the time-resolved FRET (trFRET) methodology. (A) In the U state, the donor (blue square, D) and acceptor (red triangle, A) are far apart, resulting in little or no FRET, evident in terms of the high fluorescence intensity of the donor fluorophore. In the N state, both fluorophores are closely spaced, resulting in a decrease of donor fluorescence intensity and in an increase of acceptor fluorescence intensity, as compared to the U state. (B) The extent of FRET depends upon the distance between the two fluorophores, their molecular properties, and the solvent environments. FRET efficiency has a sigmoidal dependence on the interfluorophore distance (R_{DA}) given by the Förster equation. The Förster radius, R_0 is the distance at which the FRET efficiency is 0.5, and it depends upon the properties of the fluorophores and the solvent. (C) Time-resolved fluorescence measurements involve the determination of the fluorescence intensity decay curve, represented as a plot of photon counts versus fluorescence lifetime, of the donor fluorophore. The fluorescence intensity (shown in panel A) is directly proportional to the fluorescence lifetime (τ) of a fluorophore. The resultant multiexponential decay curve can be analyzed using appropriate mathematical models for extracting information about the different fluorescence lifetimes that would correspond to populations with differential extents of FRET or R_{DA}.

the concept of foldons, secondary structural units that assemble in a unique sequential manner to form the tertiary structure[36,37] during folding.

(2) *Nucleation and Nucleation-condensation model:* In nucleation models, a folding nucleus forms either locally in a contiguous segment of sequence,[38] possibly because of hydrophobic interactions,[39] or by the diffusive condensation of specific adjacent segments of the polypeptide chain.[40] In either case, the nucleus would form in the rate limiting step, and then the rest of the structure develops around the nucleus akin to the crystallization process. The folding nucleus is expected to be present in the transition state of folding. The nucleation-condensation model suggests the formation of an unstable diffuse nucleus in the transition state, whose structure is stabilized in the subsequent rate limiting step of concerted condensation of a specific secondary and tertiary structure.[41,42] It has been suggested that the transition state of folding is an extended nucleus which is an ensemble of structures possessing some native secondary structure but disrupted native packing interactions.[43,44] Computer simulations not only have played a pivotal role in the evolution of the nucleation mechanism[45−48] but also have provided additional structural characterization of transition states.[49,50]

(3) *Hydrophobic collapse model:* This model suggests that the nonspecific clustering of hydrophobic residues cluster together due to their solvent exclusion property, resulting in a rapid initial collapse of the polypeptide chain.[51−53] The collapsed state offers a facilitated conformational search for the native-like secondary and tertiary structure formation, due to its smaller conformational space (reduced chain entropy).[54] Computer simulations, especially those utilizing lattice models, have contributed much to the understanding of how the initial chain collapse facilitates subsequent folding.[55−58] The framework model and the hydrophobic model support the existence of intermediates during (un)folding, as well as the possibility of folding via multiple pathways. On the other hand, the nucleation model presents the folding reaction as a strictly two-state process and, hence, is at odds with the existence of folding intermediates. It should be noted that the mechanisms are not exclusive. It should be noted that it is possible for the same protein to utilize one or more of these mechanisms to fold, depending on the solvent conditions.

1.3.1. Polymer Theory for Predicting the Size of the Polypeptide Chain. Basic polymer physics theory can predict different features of the folding reaction and provides insight into the role of solvent properties in defining the size of conformational ensembles populated during protein folding. The properties of a conformational ensemble of a polymer depend on three major factors:[59,60] the flexibility of the polypeptide chain, the nature and strength of various possible

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

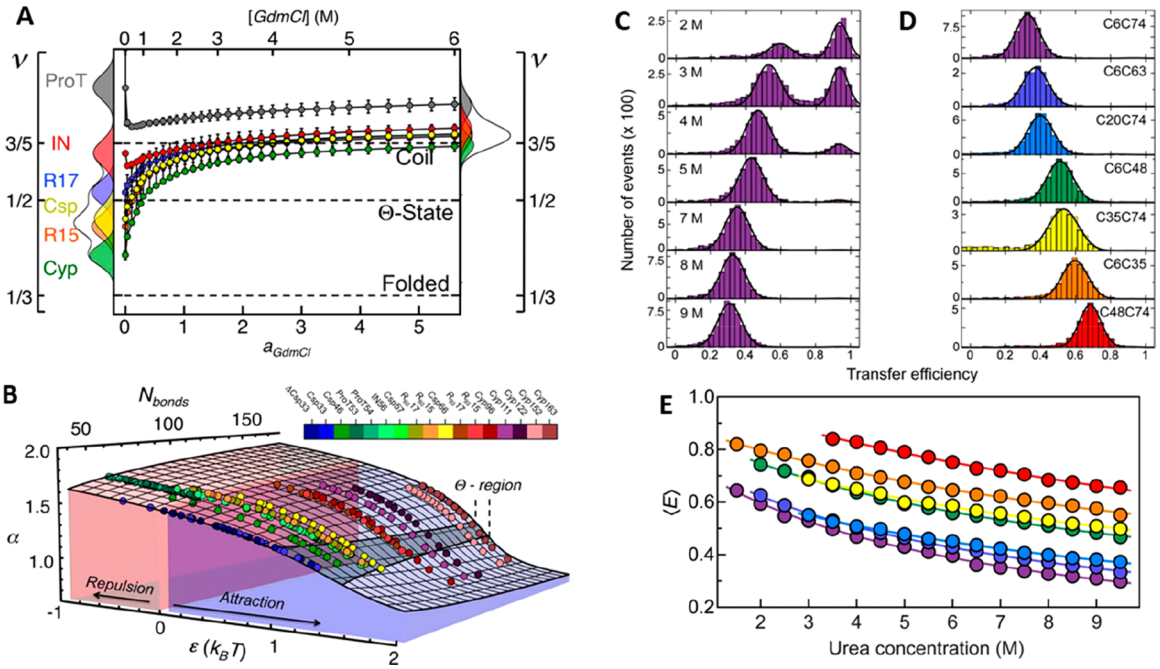


Figure 3. Characterization of the polymer-like properties of the unfolded state under different solvent conditions. Panel A shows the dependence of the scaling exponent (ν) on the activity of the denaturant (GdnHCl, in this case) for multiple unfolded proteins and their variants. At high denaturant concentrations, the value of ν approaches 3/5, affirming the random coil-like properties of the U state under denaturing conditions. Panel B depicts the phase transition surface for all the proteins/variants. The expansion factor (α) represents the change in cooperativity of the collapse transition as a function of chain length. Shaded volumes in panel B indicate the regimes of attractive ($\varepsilon > 0$) and repulsive ($\varepsilon < 0$) intrachain interaction energies. Panel C shows smFRET efficiency (E) histograms of the C6C74 variant of ubiquitin at different concentrations of urea, pH 2.5, illustrating the unfolding transition and the unfolded-state expansion. The peak at $E \approx 0.9$ arises from folded molecules and that at lower E from unfolded molecules. To determine mean transfer efficiencies, $\langle E \rangle$, peaks were fitted with Gaussian peak functions (black lines). Panel D shows the FRET E histograms at 8 M urea, pH 2.5 for all ubiquitin variants investigated, with the positions of labeled Cys residues indicated for each panel. Panel E represents the dependencies of mean transfer efficiencies of the unfolded ensemble on urea concentration for all variants (color code same as in panel D). The continuous movement in the peak position of the FRET E distributions (panel C), and the denaturant-induced expansion at all the monitored segments, confirm the solvent-dependent sensitivity of the U state dimensions and reflect the U state heterogeneity. Panels A and B have been reproduced with permission from ref 112 (Copyright 2012 National Academy of Sciences). Panels C−E have been reproduced with permission from ref 113 (Copyright 2016 National Academy of Sciences).

intrachain interactions, and the strength of interactions between the polypeptide chain and the solvent molecules. The overall rigidity of the polypeptide chain is defined by its amino acid composition. The balance between intrachain and chain−solvent interactions, which can be either attractive or repulsive or both, plays a key role in determining chain dimensions. Changes in solvent polarity or other environmental changes (pH, temperature, pressure, chemical denaturant) will also influence the conformational distribution of conformations adopted by the polypeptide chain.

To understand the evolution of conformational heterogeneity, it is important to determine how the naturally evolved heteropolymeric polypeptide chain obeys the laws of basic polymer physics[59,60] at different stages of the folding reaction.

1.4. Heterogeneity in Protein Folding

In a protein folding reaction, the U state transforms into the N state, transiting through metastable, and occasionally stable, intermediate (I) states. Each of these states is an ensemble of a large number of protein conformations and may consist of a number of substates (subensembles) that differ in their average structures and energies. The existence of conformational heterogeneity in the N state was established many decades ago,[61,62] and many studies have shown the existence of conformational substates that are in dynamic equilibrium with each other.[63−67] Protein dynamics is thought to be essential for function[68−72] and for the evolution of new function.[73,74] A protein molecule can adopt many conformations as a function of time, and a population of molecules will consist of these multiple conformations at any given time. Conformational substates can be distinguished from each other based on their free energy differences, which determine their relative probabilities of occurrence, and their time scales of interconversion.

The free energy of any conformational substate for a given protein is dependent upon the environment: it is defined by the solvent quality, temperature, and pressure. *In vivo*, proteins may experience different kinds of environments, which in turn affect the population distributions of various conformational substates.[75−77] To function, proteins may transiently populate "non-native" conformations.[71,78,79] Non-native conformations typically have higher free energies and are thus populated to very small extents. Studies capable of resolving heterogeneity are able to characterize these dynamic conformations.[2,66,70,79−84] Information about these sparsely populated substates is generally not revealed by the use of low-resolution ensemble-averaging probes. Over the past several decades, protein folding mechanisms have been investigated both by computation and by experiment.[2,7,9,10,19,20,22,37,44,47,48,75,85−97] What has only recently begun to be explored, with the availability of high-resolution probes, such as time-resolved (tr) FRET (Figure 2),[98,99] is the heterogeneity, both temporal and structural, associated with the folding transition.[2,7,10,65,100,101] Under-

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

standing the role of conformational heterogeneity at various stages of (un)folding is of utmost importance for deciphering the folding mechanism correctly and unambiguously.[2]

The conformational heterogeneity needs to be resolved first. The following questions need to be addressed: How many distinct conformational ensembles are populated at different stages of protein folding? How heterogeneous are the conformational distributions of the U state, the intermediates, and the N state? Does a protein fold via a single-defined pathway or via multiple routes? What is the time scale of exchange between different conformational substates that might be present in the U state, the intermediates, and the N state? How fast is the conversion between substates as compared to the transitions between different conformational ensembles? What are the nature and heights of the activation energy barriers slowing down different structure-forming steps? Is a protein (un)folding reaction two-state (fully cooperative), multistate (limited cooperativity), or continuous (noncooperative) in nature, wholly, or in parts? How does cooperativity vary across different structural regions of a protein? How cooperative is the initial polypeptide chain collapse?

2. RESOLVING STRUCTURAL HETEROGENEITY AT DIFFERENT STAGES OF PROTEIN FOLDING

2.1. The Unfolded State

Folding begins from the unfolded state. Hence, a comprehensive and detailed understanding of the size, structure, and dynamics of the U state is critical for gaining mechanistic insight into the folding reaction.[2,20,100,102−104]

2.1.1. Size of the Unfolded State. A completely unstructured U state, which is devoid of specific intrachain interactions, is expected to have size-scaling properties akin to those of homopolymers.[59,60,104−106] For a random coil-like polymer chain (of finite length) in dilute solutions, the R_g of the ensemble is expected to increase exponentially with an increase in chain length ($R_g \propto N^{\nu}$),[59] with a value of 0.6 for the Flory exponent, ν, obtained using mean field theory. A more precise value of 0.588 for ν has been derived using renormalization group theory.[107]

Extensive experimental measurements of the size of the U state ensemble have been made with the use of different spectroscopic probes, especially small-angle X-ray scattering (SAXS) measurements.[103,108−111] An estimate of the R_g obtained for chemically denatured, cross-link-free, and prosthetic group-free proteins ranging in length from ∼10 to ∼550 amino acid residues provided strong experimental proof (ν = 0.598 ± 0.028) for the random coil-like nature of the unfolded polymeric chain.[103] These results were corroborated later with the use of other spectroscopic measurements.[112−114] In a single molecule (sm) FRET-based study[112] of the U state of various globular proteins, as well as of intrinsically disordered proteins (IDPs), it was shown that the scaling exponent for all the proteins converged to a value of 0.62 ± 0.03 at high denaturant concentrations, indicating that chemical denaturants are good solvents (Figure 3). Recently, another comprehensive study probing the dimensions of the U state of ubiquitin using NMR, SAXS, and smFRET yielded similar values for ν (NMR + SAXS: ν = 0.61 ± 0.03, FRET: ν = 0.60 ± 0.03), reiterating that the U state ensemble at a high denaturant concentration behaves like a random coil-like polymer chain in good solvent.[113] Interestingly, chemically denatured proteins have been shown to display random coil-like behavior even when a nonrandom structure is present.[103,113,115] Measurements using the trFRET methodology to probe intrasegmental distances as a function of sequence separation, for NTL9[116] and single-chain monellin (MNEI),[117] also confirmed the homopolymer and random coil-like behavior of the U state under denaturing conditions.

The hydration of individual amino acid residues in the polypeptide chain appears to play a critical role in modulating the size of the U state, which is described by universal scaling laws. Since the free energy of hydration is dependent on temperature, it is not surprising that the size of the U state too is dependent on temperature. Both in the absence and presence of chemical denaturant, an unfolded protein chain becomes more compact upon an increase in temperature[118−120] but only up to a certain temperature above which the protein either undergoes no further significant compaction[121] or may even start expanding.[122] Remarkably, for a mutant variant of CTL-9, the temperature above which no further compaction occurs, corresponds to the temperature of maximum stability.[121] Not surprisingly, the hydrophobic effect,[123] through its dependence on temperature, seems to play an important role in determining the size of the unfolded protein. Computer simulations have also shown that the size in chemical denaturant is smaller at a lower temperature[124,125] and that solvent quality is altered by a change in temperature.[126]

2.1.2. Conformational Dynamics within the Unfolded State. In an ideal random coil state, all possible conformations would be sampled on a very fast time scale, much faster than the rates for subsequent folding events. Folding from such a state would occur along a single folding pathway.[2,127] However, the U state often comprises subpopulations of site-specifically differing conformations that interconvert on a time scale similar to, or slower than, the subsequent folding events.[2,65,83,128−130] Slow conformational changes in the U state ensemble, followed by relatively faster folding steps, result in heterogeneity at the early stages of protein folding.[2,127,131] The rate constants for conformational dynamics within the U state ensemble have been measured for several proteins and were found to vary by almost 10 orders of magnitude.[65,92,132−140] A representative but not an exhaustive list of examples can be categorized into three categories:

2.1.2.1. Very Fast Dynamics (100 ps to 100 ns). The loop closure kinetics in the U state of the Engrailed homeodomain protein, measured using photoinduced electron transfer-fluorescence correlation spectroscopy (PET-FCS), had a time constant of 200 ns, while the corresponding folding step occurred in 2−20 μs. Thus, the U state was effectively homogeneous with respect to the folding process. Recent smFRET studies on ubiquitin[113] and several other proteins[92] have suggested that chain fluctuations in the U state at high denaturant occur in less than 100 ns.

2.1.2.2. Fast Dynamics (100 ns to 10 μs). FCS measurements have revealed U state fluctuations on the microseconds time scale in the intestinal fatty acid-binding protein, and the kinetics was sensitive to solvent conditions such as pH, ionic strength, etc.[141] The acid-denatured state of apomyoglobin also undergoes conformational fluctuations on the time scale of 3−200 μs.[142] Trp-Cys contact quenching measurements on the Protein L[143] and Acyl-CoA binding protein[144] also revealed microsecond dynamics (∼10−100 μs) in the U state.

2.1.2.3. Slow Dynamics (ms to s). Equilibrium smFRET measurements also provide a measure of the time scale of exchange between different conformations within the U state ensemble based on the broadening and exchange seen in various

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

FRET efficiency histograms.[65,145−147] smFRET studies have suggested that conformational interconversion within the U state ensemble of Protein L[148,149] occurs on a time scale of more than several milliseconds, and line confocal smFRET studies, for which the photon statistics are much better, have indicated the same for the B domain of protein A, BdpA,[150] and ubiquitin.[149] In the case of Protein L, however, a subsequent study, utilizing smFRET as well as PET, did not reveal slow dynamics in the U state ensemble.[132] Interestingly, the time scale of interconversion between different conformational substates in the U state ensemble under denaturing conditions can also be on the seconds time scale, as seen for chymotrypsin inhibitor 2,[151] RNase H[129] in smFRET measurements, and MNEI[83,117,130] in ensemble trFRET measurements. It should, however, be noted that the inference of slow dynamics from the widths of FRET efficiency histograms in smFRET experiments may not be correct.[128] It was suggested that the widths in excess of shot noise might be caused by other sources, such as nonrandom photon emission intervals resulting from triplet state formation or intensity variation across the focal volume, and not necessarily due to slow interconversion between various conformational substates.

2.1.3. Origin of Heterogeneity in the U State.

2.1.3.1. Local Heterogeneity. A major cause for U state heterogeneity is *cis*−*trans* isomerization of the peptidyl−prolyl bond.[152−157] As the free energy difference between the *cis* and *trans* isomers is very small for an X-Pro bond, significant populations of both isomers can exist at equilibrium in the U state ensemble.[152] The isomerization rate is very slow (0.01−0.001 s^{-1}) due to the presence of a large activation energy (∼20 kcal mol^{-1}),[158] significantly slower than the subsequent structure forming events. Thus, peptidyl-prolyl isomerization contributes to heterogeneity at both early[154,159] and later stages[152,153,155,160] of folding. Recent studies suggest that the isomerization status of the peptidyl−prolyl bond can also play a significant role in facilitating long-range interactions in the U state by accelerating the folding reaction along a distinct folding pathway.[161,162]

Different rotational isoforms of Trp or any other amino acid residue (usually aromatic) can also contribute to local heterogeneity in the U state.[163] The scrambling of disulfide bonds in cysteine-rich proteins containing three or more Cys residues can also be a source of heterogeneity.[164−166] This has been observed during the folding of BPTI, where U state molecules that have the wrong disulfide bonds formed take a longer route to fold than the ones which have the right disulfide bonds formed or which have no bonds formed.[167,168] In the cell, the disulfide isomerases[169] and peptidyl-prolyl isomerases[170] catalyze the isomerization reactions, reducing the apparent heterogeneity in the U state ensemble. In cytochrome *c*, the His residues in the protein interact noncovalently with the heme moiety. This interaction can occur in many ways in the U state,[80,171] while in the N state only the right set of interactions is formed. When His misligation is avoided, cytochrome *c* folds within 15 ms.[172]

2.1.3.2. Global Heterogeneity. The U state is not static but is highly dynamic and malleable. Evidence for the malleability of the U state is seen in the significant slopes of the unfolded protein baselines of the equilibrium unfolding curves of several proteins, especially when structure-monitoring probes, such as far-UV CD and FRET,[7,10,101,173,174] are used. An increase in the concentration of chemical denaturants is known to increase the dimensions of the U state.[92,101,102,113,128,175−180] This observation has been made with globular protein[101,113,114,177,179,181,182] as well as IDPs.[183−185] An increase in the number of favorable chain−solvent interactions with an increasing concentration of the chemical denaturant leads to further expansion of the polypeptide chain.[92,102,175,186−188] This observation can also be understood by assuming that the U state ensemble is a heterogeneous mixture of various conformations, differing in their extents of compaction. The addition of a denaturant would re-establish equilibrium by stabilizing and thereby increasing the relative populations of the more expanded conformations with a higher solvent-exposed surface area.

Equilibrium studies utilizing both smFRET[92,104,112,113,135,183,189] and trFRET[101,177,182,190] have shown that the intramolecular distance (R_{DA}) has a significant dependence on denaturant concentration. Line-confocal smFRET measurements yield much higher photon numbers from a single molecule than conventional smFRET measurements and, in the case of both BdpA[150] and ubiquitin,[149] have been suggestive of conformational heterogeneity within the U state.[65] Two-dimensional fluorescence lifetime correlation spectroscopy (2D-FLCS) measurements have also revealed significant heterogeneity in the U state ensemble of BdpA.[191]

Site-specific conformational preferences have also been observed in the U state ensembles of several proteins.[2,65,115,175,192−198] Different conformational substates that constitute the U state ensemble differ usually in terms of the residual structure present.[192,199] Both local and nonlocal interactions are known to stabilize the residual structure present in the U state ensemble.[149,200−203] Native-like (N-like) interactions are expected to facilitate further structure formation, resulting in fast folding to the N state. Non-native interactions have also been seen in the U state ensemble, which may retard the folding reaction.[65,82,192,199,204]

The formation of local hydrophobic clusters is also known to restrict conformational freedom within the U state.[65,149,195,204−206] In the case of lysozyme, small local hydrophobic clusters formed by two Trp residues were shown to restrict the motional freedom of the fluorophores, resulting in conformational heterogeneity.[206] Similarly, nonspecific intramolecular interactions between a fluorophore adduct and the polymer chain have been shown to result in an ensemble with heterogeneous photophysical properties.[207−210] NMR and paramagnetic relaxation enhancement-based studies have also confirmed the presence of local hydrophobic clusters and heterogeneous residual structure in the U state ensemble.[193,194,196,211] In the case of NTL9,[212] ACBP,[213] and ubiquitin,[197,198] NMR measurements identified U state heterogeneity in terms of there being N-like local secondary and tertiary interactions present, observable because of slow conformational exchange times. Recent studies utilizing hydrogen exchange NMR (HX-NMR) measurements on chemically denatured ubiquitin have revealed the presence of a N-like hydrogen-bonded residual structure in the α-helix and β-hairpin regions, implicating its role in facilitating the subsequent folding reaction.[214]

2.2. The Native State

The folded form of a protein populated under native conditions is very often compact and globular[20,215,216] but not always so.[217] The N state is defined by a specific and unique set of stabilizing intramolecular interactions.[6,218] Its average size is generally describable by the model for a heteropolymer in a poor solvent,[59,60] where intramolecular chain−chain interactions are

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

preferred over solvent−chain interactions; hence, only a modest increase in the size of the polymer chain is expected with an increase in the chain length. Structural characterization of the N state is possible with the help of high-resolution probes such as X-ray crystallography. However, this measurement only provides a static image of the protein conformation that is maximally populated. In fact, even before the first X-ray-crystal structure became available, hydrogen exchange experiments had revealed that native proteins structures are dynamic.[219] Many subsequent experiments have also revealed the dynamic and heterogeneous nature of the N state.[7,62,63,84,177,220−222] Although the interior of a folded protein is very well packed, it is still not solid-like. Instead, it appears to be like a dense fluid with enough space for conformational fluctuations.[223,224]

2.2.1. Size of the Folded State. The dimensions (segmental distances) of the folded state measured using FRET have unequivocally matched with the predictions made based upon available crystal structures, validating the usage of the FRET methodology as a probe.[101,116,177,225] The population distributions obtained for intramolecular FRET efficiencies and distances for the N states of various proteins have always been narrower than those of the corresponding U state distributions, indicating the relatively homogeneous nature of the N state.[65,99,101,135,177,182,191,226] Nonetheless, in the case of barstar,[177] a SH3 domain,[182,227] BdpA,[191] and MNEI,[101] heterogeneity in the size of the folded state has been revealed. It was shown that the folded state undergoes expansion with an increase in denaturant concentration under equilibrium conditions.

2.2.2. Conformational Dynamics within the N State. Studies utilizing the hydrogen exchange methodology were the first to reveal fluctuations associated with local and global chain motions of a protein even under native conditions.[81,219,222] The N state samples various high-energy partially unfolded states as well as the global U state.[228−230] Hydrogen−deuterium exchange studies on myoglobin and insulin were the first ones to demonstrate the dynamic nature of the folded state.[219] Native state HX measurements on RNase A revealed the presence of two slowly interconverting conformations present in the folded state at equilibrium.[231] Native state thiol-exchange measurements on MNEI[229] revealed high-energy intermediate states frequently sampled by the native state, which are otherwise inaccessible to standard kinetic measurements. PET-FCS has helped to reveal the time scale of conformational fluctuations within the N state of several proteins.[84,221,232] Recently, with the use of the PET-FCS methodology, dynamics in the N state of the aggregation-prone mouse prion protein were revealed, whose time scale was modulated under aggregation conditions.[84] An observation of glassy-like slow dynamics within the N state ensemble of MNEI[83,117] and of N-terminal domain of spider silk protein[221] suggested that the barriers within the N state ensemble can be as high as the barrier corresponding to the global unfolding transition.

2.2.3. Nature of Heterogeneity in the N State. *2.2.3.1. Local Heterogeneity.* Peptidyl−prolyl bond isomerization has been shown to introduce heterogeneity even in the N state for some proteins.[233,234] For β-2-microglobulin, the N state is heterogeneous with respect to *cis*−*trans* isomerization of Pro 32.[235,236] In the N state of staphylococcal nuclease, an equilibrium between different isomers at two Pro residues (Pro117 and Pro47) was observed.[237] These Pro switches in the N state have been shown to have functional roles in some cases[233,234] and also have been shown recently to modulate the heterogeneity of the binding-induced folding of natively unstructured proteins.[238] Local heterogeneity due to Phe conformations also has been observed in the N state of the intestinal fatty acid-binding protein, where five out of eight Phe residues were found, by NMR studies, to be in two conformations in the N state.[239]

2.2.3.2. Global Heterogeneity. For several small single-domain proteins, studies utilizing the trFRET methodology have also shown that the N state is not homogeneous and that it exists in equilibrium with multiple partially unfolded and progressively destabilized forms.[99,101,177,182] An equilibrium unfolding study of barstar[177] was among the first studies that highlighted the inherent heterogeneity present even in the N state ensemble. Later, multisite trFRET measurements revealed N state heterogeneity in equilibrium unfolding measurements of a SH3 domain[182] and MNEI.[101] N state heterogeneity has also been revealed in the villin headpiece subdomain with the help of FRET-based measurements; conformational heterogeneity was seen in the position and helical content of the C-terminal helix.[240]

smFRET studies usually have limited resolution in the shorter distance regime corresponding to N state dimensions, and hence, have not been able to detect N state heterogeneity in many cases. However, in the case of the equilibrium unfolding of a SH3 domain[227] and of BdpA,[150] N state conformational heterogeneity could be visualized. smFRET studies have also revealed fast and slow dynamics in the native states of adenylate kinase[241] and the QSOX enzyme,[242] respectively. The flexible nature of the N-terminus of the native protein in a SH3 domain,[243] as well as of the mouse prion protein,[84] has also been elegantly revealed. In the case of thioredoxin, a combined assessment of the crystal structure, NMR measurements, and studies using molecular dynamic simulations suggested that alternative conformations coexist within the N state.[244] In an equilibrium unfolding measurement of BdpA, heterogeneity in both the N and the U states was revealed by 2D-FLCS.[191] Both ensembles demonstrated a dependence of size on the denaturant concentration. The structural origin of N state heterogeneity in BdpA was identified as the fraying of the N-terminal helix away from the rest of the protein.

2.3. The Collapsed Intermediate Ensemble

The folding reaction involves significant changes in size, structure, and heterogeneity of the polypeptide chain. The collapsed intermediate ensemble, U_C, marks the beginning of the folding reaction as it represents the U state under folding conditions.[2,20,55,93,245,246] It is necessary to characterize U_C to mechanistically resolve the sequence of structural events accompanying the folding transition.

In kinetic experiments, an initial collapse in the size of the polypeptide chain is usually observed upon changing solvent conditions from being denaturing to renaturing.[8,20,53,83,104,117,188,247,248] Equilibrium smFRET experiments too have indicated that the unfolded state becomes more compact as denaturant concentration is reduced.[92] These results are supported by simulation studies.[93,116,124,249−252] A fundamental question is whether the initial polypeptide chain collapse facilitates subsequent folding.[20,65,93,111,172,253] The hydrophobic collapse model (Figure 1) suggests that a change in the solvent environment from being denaturing (less polar) to renaturing (more polar) induces a collapse (reduction in size) of the polypeptide chain due to the clustering of hydrophobic residues away from the solvent to reduce their exposure to polar

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

solvent.[20,53,215] Such a collapse would be facilitated when intrachain interactions are preferred energetically over chain−solvent interactions under folding conditions, suggesting that, under native conditions, the solvent has the attributes of a poor or theta solvent (see section 2.1). Initial chain collapse has been proposed to facilitate folding by reducing the conformational search for further specific structure formation leading to the N state.[20,215] However, it has been posited that the U state should remain expanded even under folding conditions, as the formation of nonspecific hydrophobic clusters may enhance intermolecular interactions and could lead to nonreversible aggregation of the protein.[253,254] It is therefore important to characterize the properties of the collapsed intermediate ensemble, for establishing its role in folding.

Characterization of the extent and nature of structure in the U_C ensemble is essential for identifying the driving force for collapse and its role in folding. In some cases, collapse and structure formation (folding) appeared to occur simultaneously.[108,255] Nevertheless, with the help of fast mixing devices, the occurrence of a fast chain collapse preceding specific secondary structure formation has been resolved for several proteins.[20,53,124,256−259]

2.3.1. The Time Scale of Initial Chain Collapse. The time scale of the collapse transition has been measured using either pump−probe-based relaxation methods or fast mixing devices such as the continuous flow method. The collapse of a homopolymer has been proposed to occur in two stages:[105,106] a very fast local arrangement, whose time constant (t_{c1}) is independent of the number of monomers (N), followed by a slow large-scale reorganization step, whose time constant (t_{c2}) is dependent on N ($t_{c_2} \propto N^{-3/2}$, for a random coil-like polymer chain). Measurement of intramolecular contact formation kinetics between Trp and Cys residues in long peptides has provided diffusion-limited quenching rates with the time constant of collapse approaching that expected for random coil-like behavior.[133]

The time scale for chain collapse in small proteins varies by almost three orders of magnitude, from 60 ns for BBL,[118] to 10−35 μs for RNase A,[246] cytochrome c and its nonfoldable analogues,[133,260,261] and MNEI.[247] An estimate of the activation energy barrier ($4k_bT$ to $11k_bT$) for collapse obtained from equilibrium smFRET measurements provided an estimate of 40 ns for the collapse of the cold shock protein.[128] The time scale for chain collapse and folding is controlled by the solvent viscosity as well as the internal friction of the protein.[133,247,262−264]

2.3.2. Size of U_C. The size of U_C has been determined, either with the help of small-angle scattering (SAXS) studies that provide the radius of gyration of the protein[108,111,265,266] or by using site-specific FRET measurements that provide segmental or end-to-end distances (R_{DA}).[8,83,104,259,267,268] Ensemble and smFRET-based studies on several single-domain globular proteins as well as IDPs suggest that U_C is significantly smaller than the U state.[8,20,83,124,247,252,266,269] The dependence of the size of U_C on chain length suggests that it corresponds to the theta or globular state of the polypeptide chain.[20,60,270] However, studies utilizing SAXS measurements suggest that U_C is more similar in size to the U state.[108,111,253,254,266] For the folding of protein L, a clear discrepancy between SAXS and smFRET measurements was observed.[266] It is possible that FRET labels (fluorophores) may promote collapse in an otherwise unfolded polypeptide chain,[254,271] as was observed for measurements of the size of (un)labeled poly ethylene glycol solutions.[272] However, with careful experimentation and analysis, a consistent view of polypeptide expansion (as well as contraction) with increasing (or decreasing) denaturant concentration has been established for both globular proteins as well as IDPs.[114,124,251,273]

The origin of the discrepancy between inferences drawn from SAXS and FRET measurements can be traced back to the heterogeneity present in the collapsed intermediate ensemble, which precludes a straightforward comparison between R_g and R_{DA}.[83,273−275] The conversion of R_{DA} to R_g requires assumptions about the molecular shape, which do not hold under folding conditions.[124,273,275,276]

It has recently been demonstrated for cytochrome c that dramatic changes in the shape of the protein occur during collapse.[216] Multiple values of R_g can be consistent with the same average value of R_{DA} due to the heterogeneity inherent in U_C.[83,275] Such discrepancies in the dimensions measured using SAXS and smFRET are not observed for fully unfolded protein at high denaturant concentrations due to fast fluctuations in the random coil-like U state, resulting in effectively a homogeneous ensemble.[92,113,117] Thus, the observed differences seen in the extent of collapse, between SAXS and FRET measurements, are not because of shortcomings in either of the two methods and also not because of dye artifacts. Instead, they can be attributed to the chemical heterogeneity associated with heteropolymer systems such as proteins, which leads to a decoupling between two different polymer properties, R_g and R_{DA}, which gets amplified in the absence of denaturant.[83] A comprehensive understanding of the size of the U_C ensemble can be obtained by combined measurement of R_g and R_{DA}, along with an assessment of the polymer shape.

The collapse and folding of MNEI have been studied extensively using multiple methods. Both SAXS[109,251] and FRET[83,117,162,247,259]-based studies suggested an initial collapse in the size of the protein occurring within 37 μs of the commencement of folding. On the basis of both steady-state and time-resolved FRET measurements, the extent of reduction in size from U to U_C within the first 37 μs of folding for MNEI was estimated to be ∼40%.[83,117,247] Interestingly, this varied across different structural regions of the protein.[117,162,247,259] The nonuniform collapse has also been observed in the case of *E. coli* adenylate kinase,[257] barstar,[180] a SH3 domain,[267] and ubiquitin.[124] In the case of MNEI, the helix segment undergoes only a small extent of collapse, but the parts spanning the β-sheet collapse significantly.[117] This result agrees with a recent analysis that suggests that initial chain collapse during folding occurs for all proteins but that the extent of the collapse is a function of the structure and topology of the final folded state.[93,270] The theory also suggests that proteins rich in β-sheets are more collapsible than α-helical proteins. Thus, the size of the collapsed intermediate ensemble will depend largely on the primary sequence and, hence, the structure of the protein. The size and interspacing of hydrophobic clusters along the polypeptide chain determine the strength of the force driving collapse.[93,111,253,254] A recent analysis of the sensitivity of IDP dimensions on solvent conditions, also reveals that the size of the polypeptide chain has an intricate dependence on both the sequence and the solvent quality.[277]

2.3.3. Structural Heterogeneity in U_C. U_C is conformationally heterogeneous, with some parts of the protein collapsed in some molecules and others expanded. Examples of proteins with heterogeneity observed in U_C include TIM barrel

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

protein,[156] barstar,[8,180] dihydrofolate reductase,[268] a thermophilic variant of cytochrome *c*,[278] NTL9 protein,[116,203] and MNEI.[83,117,162] Detection of a large width for the fluorescence lifetime/distance distribution of the U state under native conditions is another piece of evidence for heterogeneity in the U_C ensemble.[83,279] The observation of slow interconversion rates within the compact U_C ensemble also provides evidence for the heterogeneous nature of Uc.[280,281]

2.3.4. Specificity of the Structure of Uc. It is important to note that the structure apparent in U_C depends on the stage of the reaction when it is probed. As described earlier, collapse for several proteins occurs in <50 μs;[20] however, chain rearrangement after the initial collapse continues to result in structure evolution.[247] Thus, the structure present in the collapsed state observed at a few ms of folding has most likely evolved after the fast initial chain collapse.[20] A specifically collapsed intermediate with N-like interactions is expected to act as a productive folding intermediate in facilitating further folding.[260,282,283] On the other hand, a nonspecifically collapsed intermediate might represent merely the readjustment of the U state ensemble in response to a change in solvent properties, which may or may not facilitate further folding.[284] In fact, in some cases, initial chain collapse has been shown to retard folding by increasing the height of the folding barrier.[154,172] The observation of differential extents of contraction across different parts of the protein would suggest that the collapsed intermediate is site-specifically collapsed, as seen for barstar,[180] staphylococcal nuclease,[285] monellin,[117,247,259] and NTL9.[116] Other signatures such as the presence of an N-like secondary structure and spectral properties have also been considered to represent specific collapse, as seen for CspTm,[137] RNase A,[109,246] apomyoglobin,[286] *E. coli* DHFR,[268] and cytochrome *c*.[287]

However, in the case of some proteins, the fractional change in FRET efficiency across different segments of the protein is similar, and the U_C state was shown to be devoid of N-like spectral properties, suggesting that initial chain collapse is nonspecific as observed for the IDP alpha-synuclein,[288] the SH3 domain of PI3 kinase,[267] and staphylococcal nuclease.[285] The collapse in the case of MNEI has both specific and nonspecific components,[247,251,259] which could be identified by resolving the heterogeneity of the U_C ensemble with the help of a multisite trFRET-based folding kinetics study.[117] In the case of barstar too, specific and nonspecific components in the initial collapse could be observed during folding under very strongly native conditions.[8,180]

2.3.5. Nature of Interactions. While the formation of N-like interactions in U_C is expected to facilitate further folding to the N state,[215,289−291] even non-native interactions have been shown to form during the folding reaction.[83,202,247,292,293] Early evidence for the presence of non-native interactions during initial stages of protein folding was observed for hen egg white lysozyme[293] and β-lactoglobulin.[292] In the case of the all β protein β-lactoglobulin, the non-native helix formed in U_C at a few milliseconds of folding has been proposed to facilitate folding by enhancing nonlocal interactions.[292,294] Similarly, a non-native, on-pathway collapsed intermediate has been shown to populate within a few μs during the folding of *E. coli* RNase H.[295,296] The burst phase intermediate populated within 37 μs of the folding of MNEI also has non-native interactions which appear to prime further formation of the N state structure.[83,247] While these studies support a productive role for U_C, other studies suggested that the population of U_C retards an otherwise fast-folding process.[154,172] During the folding of MNEI, the nonspecifically collapsed unfolded-like (U-like) subensemble in the U_C ensemble has been identified as a metastable misfolded intermediate due to the large activation energy barrier associated with its further collapse and folding.[83]

It appears that initial polypeptide chain collapse is driven by a hydrophobic effect, where a change in the solvent environment initiates hydrophobic cluster formation within the protein chain. During initial chain collapse, intramolecular hydrogen bonds also form at the expense of chain-solvent hydrogen bonds. Identifying the interactions that drive initial polypeptide chain collapse is an active area of research.[20,188,269] The hydrophobic nature of U_C has been revealed in the case of various proteins, by virtue of their binding to the dye ANS, which binds to solvent-accessible hydrophobic patches.[53,180,246,247,267] However, significant secondary structural content has also been observed in the U_C of various proteins, usually at a few milliseconds of folding.[137,268,286,287] The role of hydrophobic and hydrogen-bonding interactions can be delineated by studying the temperature dependence of the reaction[297] and by site-specific mutagenesis. A recent study has proposed that hydrophobic forces play a major role in driving the folding reaction via the formation of an N-like specifically collapsed intermediate.[253]

It is important to consider the nature of the hydrophobic effect, given the role played by it in polypeptide chain collapse. The hydrophobic effect is a principal force driving protein folding.[55,298] Its origin lies in the decrease in entropy and hence the consequent increase in free energy which occurs upon the hydration of a nonpolar solute.[298] The free energy cost is found to scale linearly with the volume of the solute in the case of linear alkanes,[299] as proposed for small solutes that can be accommodated without any breakage of hydrogen bonding between water molecules.[123] This decrease is traditionally attributed to the formation of a clathrate[300] or shell[301] of water molecules around the nonpolar solute, which are hydrogen bonded to each other and/or tethered by van der Waals interactions to the solute.[301,302] The hydrophobic effect is, however, not just due to the ordering of water molecules.[303] Another possible contribution is the entropic cost of forming a cavity of appropriate size to accommodate the nonpolar solute because of the small size of water molecules.[304,305] A third possible contribution could be the entropic cost associated with the decrease in translational motion of water molecules because they are excluded from the volume occupied by cavities accommodating the solute molecules.[306] In all cases, the increase in the entropy of water molecules upon burial of nonpolar surface area is the hydrophobic interaction or force that drives polypeptide chain collapse and protein folding. The relative importance of each contribution to the hydrophobic effect is yet to be delineated, especially for the initial chain collapse reaction during folding.

3. HETEROGENEITY IN FOLDING AND UNFOLDING PATHWAYS

In the vast majority of experimental studies of protein folding, the reaction has been found to be describable as occurring along one folding pathway.[22,31,307] In most studies, however, a single experimental probe, usually fluorescence or circular dichroism, has been used to monitor folding. Such probes are ensemble-averaging, and hence, the measurement yields an averaged description of all molecules present together at any instance of time. Moreover, the reaction is monitored along one reaction coordinate corresponding to the fractional change in the spectroscopic signal. The probes typically used cannot

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

distinguish between different populations of coexisting folding molecules, for example, between the U, I, and N states. Only high-resolution probes, such as time-resolved ensemble FRET[98,99,162] as well as hydrogen exchange-mass spectrometry (HX-MS),[7,222,308−310] are able to reveal the heterogeneity inherent in protein folding reactions.

An unique pathway implies that folding occurs by a unique sequence of structural events. The assumption that folding occurs via a single pathway defined by a single transition state simplifies the analysis of kinetic experimental data, whether by φ-value analysis[311] or by hydrogen exchange NMR[23,33,36,312] or HX-MS.[81,230,313,314] In particular, it has led to the foldon hypothesis, which posits that the native protein structure is assembled by the sequential assembly of specific secondary structural units in a defined order.[37]

The application of the concepts of statistical mechanics led to a very different view of the folding reaction.[315] Energy landscape theory suggests that each protein molecule folds randomly driven by thermal fluctuations, with its own unique trajectory, traversing a multidimensional free energy landscape in which an energy bias favors the native state.[86,90,96,316−319] The resulting picture is that of a folding funnel (Figure 1B) depicting the population of folding molecules using a multitude of routes to progress toward the N state. A folding pathway would correspond to a large set of folding trajectories averaged or "coarse-grained" according to some experimental metric, and different folding pathways would correspond to different sequences of structural events, averaged over different sets of trajectories. Different folding paths could be preferred under different conditions based upon the relative energies of the U state and the other states (TSs and intermediates) populated, according to their Boltzmann distributions, along the folding pathways.[90,320,321]

The protein folding reaction involves a large reduction in the conformational entropy of the polymer chain.[2,96] Given that the U state is highly heterogeneous, different conformational subensembles within the U ensemble may develop structure differently, especially if interconversion between its different subensembles is slow compared to subsequent folding.[127,131] The availability of multiple routes for the correct folding of a protein provides robustness to the folding process under a variety of environmental conditions.[2,322] It is important to identify and structurally resolve multiple pathways of folding to establish their roles in the folding.

There is significant experimental evidence that many proteins fold and unfold via parallel folding routes.[2,90,95,174,297,323−334] The folding of hen-lysozyme,[330,333] the folding[331,334] and unfolding[297] of barstar, and the folding[83,162,174,325] and unfolding[174,335] of MNEI all occur via more than one pathway. (De)stabilization of intermediates populated on parallel pathways upon a change in experimental conditions has been seen for lysozyme,[333] maltose-binding protein,[336] MNEI,[335] and ubiquitin.[329] Characterization of the difference in the structural properties of folding and unfolding transition states has helped in distinguishing parallel pathways.[297,337−340] Switching from one pathway to another has also been observed upon a change in folding conditions.[323,335,337,338,341] A recent study of the mouse prion protein[338] demonstrated that pathways utilized during folding and unfolding can be different, as their utilization strongly depends upon the starting conditions.[174,342,343] Another recent study, utilizing time-resolved SAXS-based measurements of the folding of equine heart cytochrome *c*, has also revealed the presence of multiple folding pathways arising from a highly heterogeneous U state.[131]

Molecular dynamics simulations too have revealed the existence of multiple folding pathways.[45,344−347] They have provided structural detail about how folding pathways may differ from each other.[143,251,348−352] Simulations have been particularly useful when used in conjunction with experimental characterization of folding reactions.[143,353−358] It is worth noting that while some simulations have suggested single pathway folding because of the reaction coordinate chosen,[34,359,360] reanalysis of the same simulation data using Markov state models revealed multiple folding pathways.[361]

While it is not easy to show, by experiment, the existence of multiple pathways for folding and unfolding, it is even more difficult to distinguish between the pathways on the basis of how they differ in the sequence of structural events. In the case of several proteins, the experimental data suggested that the pathways change upon a change in (un)folding conditions[297,335,338,342] as well upon mutation.[323,341,362,363] These studies indicated that the sequence of structural events on competing pathways must be different, but they could not provide any detailed structural delineation of the differences. In the case of both the early[364] and late[226] folding intermediates of barstar, changes in the structures of the folding intermediates upon changing environmental conditions were consistent with parallel folding pathways being structurally distinct.

Recent experiments utilizing hydrogen exchange-mass spectrometry (HX-MS) measurements have suggested a hierarchical sequence of structural events describing the folding of RNase H and cytochrome *c* via cooperative contiguous structural units called foldons.[36,313,365] The results were interpreted to suggest that folding occurs via a single-defined pathway and not by multiple pathways as proposed by energy landscape theory.[37,90,307,366] The hypothesis that folding must occur via a single-defined pathway led to the proposition that the heterogeneity observed during the folding of several proteins might be due to the presence of off-pathway intermediates, referred to as optional errors during folding.[36,37,367] It is, however, important to note that although HX-based measurements suggest a specific sequence of structure-forming events, they do not rule out folding via parallel pathways. Instead, the observation that the population evolution kinetics seen during the folding of RNase H and cytochrome *c* are multiexponential (even after correcting for optional errors) cannot be explained without invoking multiple pathways.[36,313,365] In fact, native state HX-MS studies on the turkey ovomucoid third domain[368] were among the first ones to reveal multiple unfolding pathways. Also, HX-MS studies of the unfolding of MNEI[369] and a SH3 domain[370] suggested that the unfolding in the presence of chemical denaturants involved independent and parallel pathways. RNase A was also shown, from native-state HX-NMR studies,[231] to unfold via competing parallel pathways.

A major recent development in the field of protein folding has been the determination of transition path times and transition path time distributions, using smFRET[371−373] as well by using single molecule force spectroscopy[374−376] measurements. Remarkably, the transition path times appear not to depend on the actual folding times ("waiting" times). These single molecule methods have the potential to directly measure different properties of the (un)folding transition. The distribution of transition path times could give insight into the extent of roughness in the energy landscape, and with further refinement it should become possible to determine how the

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

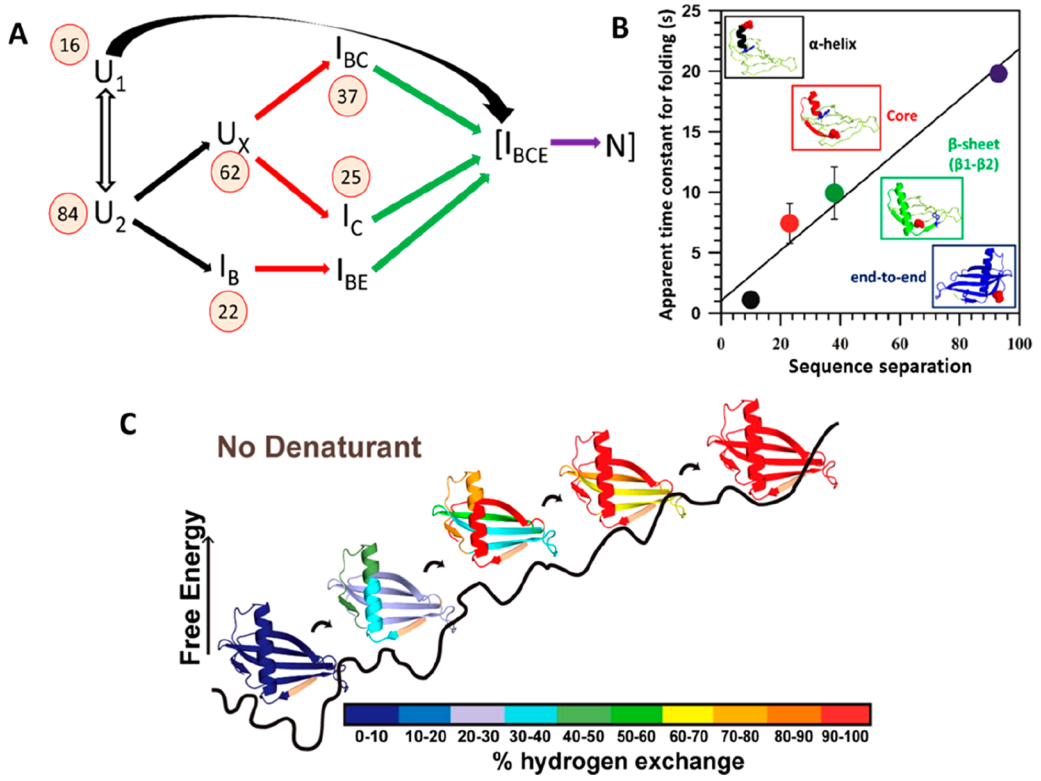


Figure 4. Structural mechanism for the folding of monellin (MNEI) derived from trFRET and HX-MS measurements. Panels A and B represent the sequence of structural events accompanying the folding of MNEI, adapted from a multisite trFRET study.[162] Reprinted with permission from ref 162. Copyright 2021 American Chemical Society. The different structural parts monitored in the study were α-helix (segment H); core (segment C); β-sheet (segment B); end-to-end (segment E). (A) The multipathway folding mechanism of MNEI. The black, red, green, and violet arrows represent the unobservable (over within 100 ms), fast, slow, and very slow kinetic phases of the folding reaction, respectively. The U state ensemble (U_1 and U_2) gives rise to three subensembles: U_X (expanded at all the monitored structural segments), I_B (collapsed at segment B), and I_{BCE} (collapsed at segments B, C, and E). I_{BCE} continues to evolve gradually to form the N state. The numbers in the circles represent the percentages of molecules following a given folding route. (B) The dependence of the population-averaged apparent time constant on sequence separation for the different FRET pairs. Different structural segments are shown in different colors corresponding to the various FRET pairs; all the residues spanned by a given FRET pair are shown with one color. The positions of the FRET donor and acceptor are shown as a blue ring and as a red sphere, respectively. (C) Sequential loss of secondary structure during transient unfolding of MNEI under native conditions. The percentages of exchange out from different sequence segments of fully deuterated native protein, which represents the extent of unfolding at different times of exchange in 0 M GdnHCl, have been mapped onto the protein structure. Both the tr-FRET and HX-MS measurements suggest that helix formation precedes the formation of core; i.e., the helix-β-sheet interface and the β2-β3 strands form early during folding. In addition, the two measurements also provide complementary information about structure formation in the parts that could be monitored exclusively in only one of the two studies. Panel C is reproduced with permission from ref 230. Copyright 2016 American Chemical Society.

folding trajectories of individual protein molecules differ from each in their sequences of structural events.[372,373,377] Single-molecule force spectroscopy measurements have also provided strong evidence for the existence of multiple pathways in the (un)folding of single-domain globular proteins,[339,378−383] as well as in the misfolding of the prion protein.[374,384] With the help of single-molecule studies, heterogeneity in the binding-induced folding reactions of IDPs has also been revealed.[238,385]

An investigation of the folding kinetics of MNEI using multisite trFRET measurements has also revealed structurally distinct parallel pathways of folding.[162] Because of the availability of site-specific population evolution kinetics data, it became possible to deduce the sequence of structure-forming events on the parallel pathways (Figure 4). Importantly, it was revealed that structure can evolve independently in different molecules, in multiple steps, involving different sequences of structure-forming events. Interestingly, the pathway-averaged sequence of structure formation suggested hierarchical accumulation of the structure, where the formation of local contacts preceded the formation of nonlocal contacts and global structure evolution. The sequence of structure formation of MNEI obtained from trFRET measurements[162] was consistent with the previously determined sequence of structure dissolution in native conditions, measured by HX-MS.[230] This highlighted the consistent and complementary nature of the two probes (Figure 4). Importantly, this study revealed that apparently hierarchical structure formation does not exclude parallel pathways of folding,[28,322,367,386] contrary to what had been suggested by the single-defined pathway model.[37]

4. UNDERSTANDING COOPERATIVITY BY RESOLVING CONFORMATIONAL HETEROGENEITY

It is important to understand the nature of the energy barriers separating the U and N states. If the two states are separated by multiple small free energy barriers, whose height is on the order of thermal energy ($\sim 1-3\ k_bT$), then the (un)folding process would proceed gradually through a continuum of intermediates. Structure formation (or dissolution) would occur continuously, and different parts of the protein would (un)fold independently or noncooperatively.[230] If, however, there is a large ($>3\ k_bT$)

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

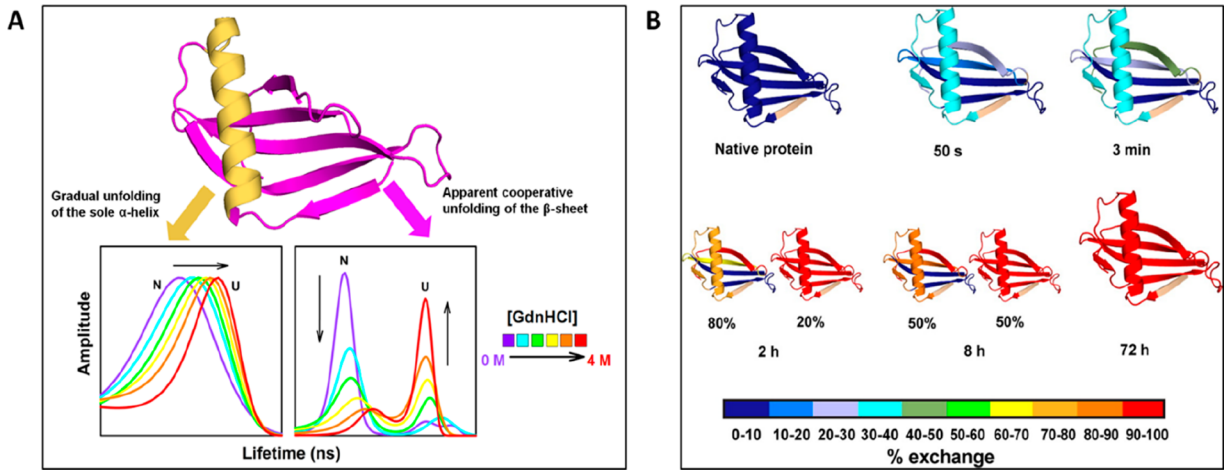


Figure 5. Site-specific resolution of cooperativity during the equilibrium and kinetic unfolding of monellin (MNEI). (A) The cooperativity of folding/unfolding of MNEI was resolved using multisite equilibrium time-resolved FRET measurements. It was shown that equilibrium unfolding is not only heterogeneous and not two-state but that the degree of noncooperativity differs between the sole α-helix and different parts of the β-sheet. While the unfolding of the β-sheet involved both cooperative and noncooperative changes, the α-helix unfolded entirely in a noncooperative manner. (B) Sequential loss of secondary structure during the transient unfolding of MNEI in 1 M GdnHCl probed using HX-MS measurements. The percentages of exchange out from different sequence segments of the native protein, at different times of exchange in 1 M GdnHCl, have been mapped onto the protein structure. Unfolding involves both continuous and cooperative structural changes. The structures of the two populations which differ in the extent of exchange at the cooperative unit at 2 and 8 h of exchange are shown. The cooperativity during unfolding also arises in the β-sheet region, in the β2 and β3 strands. Panel A has been reproduced with permission from ref 101. Copyright 2018 Royal Society of Chemistry. Panel B has been reproduced with permission from ref 230. Copyright 2016 American Chemical Society.

activation (free) energy barrier separating the U and N states, it would result in a two-state (un)folding process.[387] This would be an all-or-none cooperative process, where different parts of the protein simultaneously undergo structural change in a single step crossing the large activation energy barrier. An intermittent scenario, in which a finite number of intermediates are populated between the U and N states during (un)folding, is classified as limited cooperativity. In this case, individual steps may be cooperative, but the overall reaction is noncooperative. A recent review[7] extensively covers various aspects of protein folding cooperativity.

4.1. Cooperativity in Equilibrium Experiments

In equilibrium measurements of protein folding, studies using ensemble-averaging probes can typically identify and characterize only those conformational ensembles that are populated to a significant (>5−10%) extent. Subpopulations of molecules that are sparsely populated, usually remain undetected. Not surprisingly then, many proteins appear to undergo two-state unfolding in the presence of commonly used chemical denaturants.[387] It is only with the combined use of multiple ensemble-averaging probes,[310,388−391] or with the help of population-sensitive probes,[101,150,177,182,392] that it becomes possible to correctly determine the cooperativity of equilibrium unfolding transitions.[7,10,393]

Whether or not an equilibrium unfolding transition occurs gradually will be revealed only if the experiments are designed appropriately. The most appropriate and unambiguous probe for resolving the cooperativity of a folding reaction is the measurement of population distributions. A site-specific trFRET study on barstar had revealed continuous structural changes accompanying the equilibrium unfolding of barstar, both in the N and U states.[177] Multisite trFRET measurements on the SH3 domain of PI3 kinase[182] and MNEI[101] also revealed non-cooperative equilibrium unfolding transitions. Several smFRET-based measurements have shown that the U states of globular proteins[113,129,135,144,149,394,395] as well as of IDPs[92,112,183−185] can undergo continuous expansion with increasing denaturant concentration. Equilibrium unfolding studies of a SH3 domain[227] and BdpA,[150] probed by smFRET, suggested that both the N state and U state ensembles undergo a continuous expansion with increasing denaturant concentration.

Recent multisite trFRET measurements of the equilibrium unfolding of MNEI[101] have also revealed segment-specific variations in the cooperativity of unfolding (Figure 5). It was shown that both cooperative and continuous structural transitions accompanied the unfolding reaction. The extent of cooperativity varied across different structural segments of the proteins. The segments spanning the β-sheet were involved in both cooperative and continuous structural changes, while segments spanning the α-helix underwent only continuous noncooperative equilibrium unfolding. For several proteins now, it has been observed that the structures (and sizes) of both the N and U states are sensitive to the concentration of the chemical denaturant.[101,149,177,182,191,227] Both the N and U states appear to undergo barrier-less continuous expansion with increasing denaturant concentration.

In NMR-monitored unfolding transitions, the observation of different midpoints for the unfolding transitions of different amino acid residues in a protein has also served as good evidence for a multistate unfolding reaction.[389,391,396,397] Recent studies have demonstrated that barrier-less conformational changes in the Cnu protein support its functional requirements of being a molecular rheostat.[71]

4.2. Cooperativity in Kinetic Experiments

Kinetic studies probing the cooperativity of folding and unfolding reactions can be more informative, as they probe the sequence of structural events accompanying the (un)folding transitions. The identification and characterization of high-energy (low-populated) intermediates are more feasible in kinetic studies, which can identify transiently populated species.

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

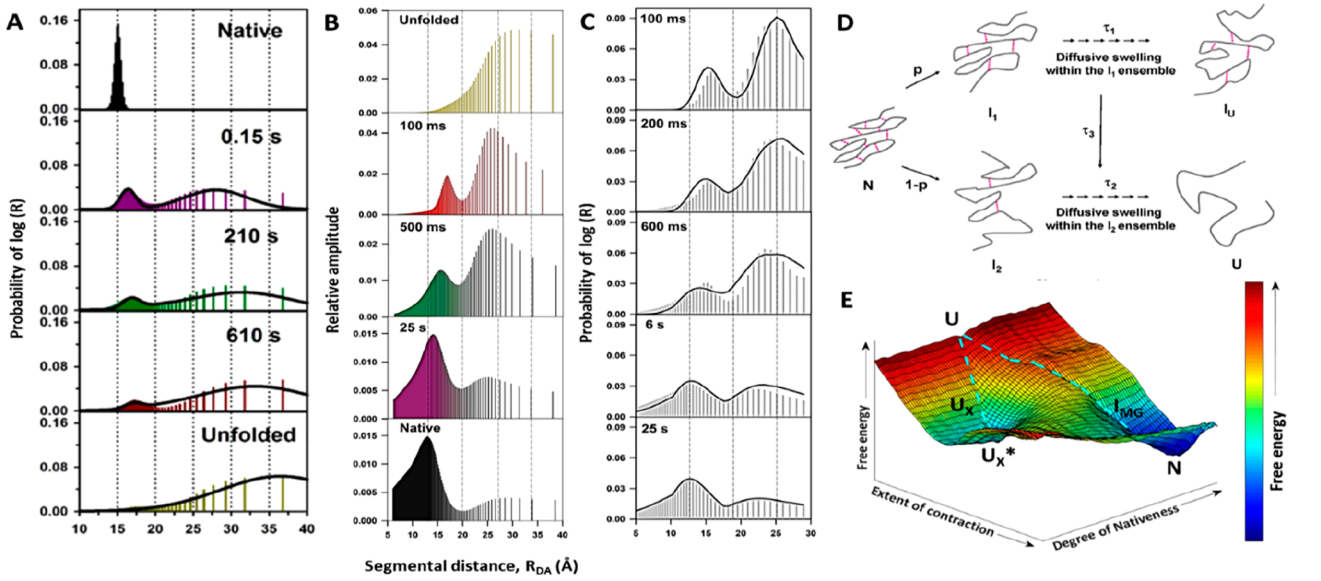


Figure 6. Barrier-less continuous structural changes accompany both the folding and the unfolding of monellin (MNEI). Panels A−C represent the evolution of intramolecular distance distributions (derived using trFRET measurements) during the unfolding and folding of MNEI. Different time points of the reaction are shown in each panel. During both folding and unfolding, continuous changes in the intramolecular distance take place, as seen in the shift of the distance distributions as a function of reaction time. The black continuous curves in panel A represent the distance distributions, derived using a modified version of the Rouse model of polymer physics (Jha et al. 2009),[130] at the indicated times of unfolding. Panel C shows a comparison of experimental (gray vertical bars) and simulated data (black solid line) obtained using a coarse-grained Markov evolution model.[83] Panel D shows the slow diffusive swelling of a Rouse-like chain with some additional noncovalent, intramolecular interactions, which has been used for modelling the continuous swelling observed during unfolding. Panel E depicts a schematic free energy landscape for describing the collapse and folding of MNEI. The free energy of folding is plotted as a function of two structural parameters, the extent of contraction, and the degree of nativeness. All the three axes have arbitrary units. "U" is the unfolded, "UX to UX*" constitute the U state sub-ensemble, and "IMG to N" constitute the N state sub-ensemble. The major pathways for folding are shown by dashed lines. Panels A and D are reproduced with permission from Jha et al. 2009.[130] Copyright 2009 National Academy of Sciences. Panels B, C, and E are reproduced, with modification, with permission from ref 83. Copyright 2019 Elsevier.

Multiple proteins have been shown to fold and unfold via discrete intermediates demonstrating limited cooperativity in (un)folding.[10,22,23,174,191,226,229,297,334,398−400] The existence of kinetic molten globule (MG) intermediates, both dry molten globules[401−404] and wet molten globules,[53,83,180,246,247,267,282,405] is another prime example of limited cooperativity.

Early events in protein folding may have reduced cooperativity.[83,109,247,259,267,294] Observations such as burst phase amplitudes having a probe dependence[200,406] or having nonsigmoidal dependences on denaturant concentration,[180,247,267,284,407] rate constants being independent of denaturant concentration,[8,20] and activation energies for collapse being small[8,173,247] all suggest that the sub-millisecond collapse/folding reactions are noncooperative. However, the observation that the kinetics of the initial chain collapse reaction is invariably exponential in nature was considered to be evidence for barrier-limited collapse.[260] However, even a noncooperative gradual transition can occur with exponential kinetics.[408,409] It was suggested that a single dominant barrier to collapse might originate from electrostatic repulsion in the U state, which could prevent collapse, and that sufficient compaction is needed for the hydrophobic effect to overcome the repulsive forces.[410] It was also proposed that the barrier to collapse might represent an uphill search for the formation of an N-like specific transition state, which can nucleate further folding in a downhill manner.[407]

When folding kinetics was monitored using population-sensitive smFRET measurements on CspTm[128] as well as by using trFRET measurements on cytochrome c[110] and MNEI,[83] definitive evidence for the barrier-limited nature of the initial hydrophobic collapse reaction was obtained. Interestingly, the cooperativity of initial collapse varied depending upon the structural part of the protein that was monitored.[162] It was observed that the segment corresponding to the α-helix underwent both collapse and folding in a barrier-less continuous manner, while the segments spanning other regions such as β-sheet or loops underwent larger barrier-limited changes in size.[117,162] Initial chain collapse was found to also involve barrier-less structural changes in both the N state and U state ensembles occurring on a time scale corresponding to glass-like transitions (Figure 6).[83,117] Similar continuous structural changes had been seen earlier in a trFRET-monitored unfolding kinetics study of MNEI[130] (Figure 6A). Interestingly, the barrier-less continuous changes, observed during both folding[83] and unfolding,[130] could be described well using simple models of polymer physics (Figure 6D). Recent HX-MS-based studies probing the kinetics of the (un)folding of MNEI[230] have revealed that folding and unfolding are completely diffusive and barrier-less in the absence of denaturants (Figure 4C). This study established that even the secondary structural elements of the protein can form and dissolve gradually in a barrier-less manner and not necessarily in a foldon-dependent manner via a single-defined pathway model.[37]

4.3. Tuning Cooperativity of Protein Folding Reaction

Cooperativity can be induced in proteins either by changing temperature,[411,412] solvent conditions,[369,370] or mutations.[413,414] A two-state transition can be converted to three-state by stabilizing an on-pathway intermediate, either by

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

changing solvent conditions[310,364] or by mutations.[415] Switching from one (un)folding pathway to another can also be achieved by modulating the stability of the TS on competing pathways.[297,416] More importantly, the transition from a cooperative two-state to fully noncooperative gradual folding can be achieved by perturbing the relative stabilities of the N and U states.[369,370,412,414] It has recently been shown, also at the level of different secondary structural elements, that the addition of denaturant stabilizes the U state of MNEI in such a manner that the noncooperative unfolding transition becomes cooperative.[230] In that study, it was shown that the $\beta2-\beta3$ region of the N state underwent cooperative (two-state) unfolding in the presence of chemical denaturants (Figure 5B). This observation suggested that the $\beta2-\beta3$ region would first fold, forming the folding nucleus. Similar modulation of site-specific cooperativity by the addition of chemical denaturant has also been observed for the SH3 domain of PI3 kinase.[370]

5. PHYSICOCHEMICAL BASIS FOR HETEROGENEITY AND NONCOOPERATIVITY

The protein folding reaction is modulated by two main independent variables: properties of the polypeptide chain (chain length and its sequence) and solution conditions. Multiple kinds of interactions exist within the polypeptide chain, between the chain and the solvent, and between solvent molecules. The trade-off between different kinds of opposing interactions may result in the stabilization of multiple conformations and lead to heterogeneity in the reaction.[5,6,417] It is the imperfect compensation of enthalpy and entropy changes during folding which gives rise to the free energy barriers.[2,6,7,69,393,413,418] In the case of the initial chain collapse, it is the balance between intrachain and chain-solvent interactions which decides the size and heterogeneity of the resultant conformational ensemble.[20,65,188] During initial chain collapse, it is also expected that if hydrogen bonding and hydrophobic interactions act consistently, then the collapsed intermediate is likely to be homogeneous;[6,65,215] otherwise, a trade-off between the two might result in a heterogeneous collapsed intermediate ensemble.[5]

The role of trade-offs between various kinds of interactions during folding was highlighted in the consistency principle.[6] The interplay between local and nonlocal interactions is likely to play a major role in determining the cooperativity of the protein folding transition.[5,91,419] While local interactions are kinetically more favorable as their formation involves lower entropic costs, the extent of structural stabilization provided by nonlocal interactions[420] via extensive packing benefits makes them thermodynamically preferred. Local interactions are expected to form faster and in a noncooperative manner, while the formation of nonlocal interactions is likely to involve the crossing of a high activation energy barrier and, hence, would be slow and cooperative.[229] Recent multisite trFRET based kinetic studies of the folding of MNEI support such a view.[162] They show that the overall sequence of folding involves the fast formation of local contacts, followed by the later formation of nonlocal contacts. Equilibrium[101] and kinetic[162] measurements of the folding of the helix in MNEI suggest that the helix, which is stabilized predominantly by local interactions, folds in a completely noncooperative, gradual manner. This has also been observed for peptides undergoing helix−coil transitions.[421,422] Moreover, the segments spanning the β-sheet (stabilized mainly by nonlocal interactions) demonstrate cooperative coil-to-globule transitions,[101,162] and the extent of cooperativity differs across the segments being monitored.

The effect of chain length on folding cooperativity depends on the interplay between packing and energetic requirements.[7,91] While an increase in the number of amino acid residues would better satisfy packing requirements, the probability of trade-offs between various kinds of interactions also increases with increasing chain length.[423] The multiplicity of possible noncovalent interactions and their spatial and temporal occurrence is the basis of the heterogeneity and noncooperativity prevalent in folding reactions.[2,7,20,35]

6. SUMMARY

Concepts from polymer physics and energy landscape theory have been important for gaining physical insight into the origin of conformational heterogeneity at different stages of the folding reaction. The use of experimental methods with the capability to resolve distinct conformational ensembles coexisting in a population has enabled a detailed understanding to be obtained on various aspects of the free energy landscape of folding. The U state under folding conditions, from which the folding reaction commences, is site-specifically heterogeneous, which leads to folding occurring via multiple pathways. Conformational heterogeneity, manifested in the form of multiple folding pathways, could be essential for folding *in vivo*, as it provides robustness to the folding process. The extent of cooperativity of folding and unfolding reactions varies in a site-specific manner between different segments of the protein and is also sensitive to solvent conditions. Initial chain collapse involves both cooperative and noncooperative transitions, and both types of transitions also occur at later times of folding for one or more subensembles of molecules. The heterogeneity seen in protein folding and unfolding reactions arises because of trade-offs between various kinds of physicochemical interactions and also because of functional or evolutionary constraints.

7. OUTLOOK

The use of high-resolution population-sensitive probes has led to important insights being obtained into the mechanistic aspects of the protein folding reaction. Resolving heterogeneity of the collapsed intermediate ensemble has provided fundamental insights into the properties of the unfolded state under folding conditions. Future studies utilizing population-sensitive probes to study structural and polymer properties of the unfolded state under a variety of solvent conditions, including cell-like crowded conditions, will lead to an understanding of how the trade-off between chain−chain and chain−solvent interactions varies with solvent type and affects conformational properties of the unfolded state as well as the energy barrier connecting the U and U_C states. The origin of heterogeneity remains poorly understood. It will be important to determine how different parts of the polypeptide chain, backbone, and side-chain units interact with each other and with the solvent. It is expected that studies involving variations in protein sequence, topology, and solvent conditions will provide insight into possible trade-offs between various physicochemical interactions that lead to heterogeneity at different stages of (un)folding reactions. Such studies will also lead to an understanding of the structural and energetic origins of reduced cooperativity in protein folding reactions. Evidence for structurally distinct and independent pathways for protein folding is still limited. The combined use of complementary high-resolution methods such

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

as time-resolved SAXS (for probing R_g and shape), multisite smFRET or trFRET (for probing tertiary structure), and HX-MS (for probing secondary structure) will enable the construction of real-time comprehensive structural images of protein molecules as they fold. It would be interesting to study the structural mechanism of folding of evolutionarily related proteins to examine the role of evolution in determining the flux and the sequence of structure-forming events across different pathways of folding. An understanding of the folding cooperativity for proteins inside the cell has remained largely unexplored. Studies resolving conformational heterogeneity need to be extended beyond the regime of self-assembly of polypeptide chains *in vitro*, for understanding folding inside the complex environment inside cells. It is likely that protein folding pathways inside cells will be modulated by the specific environment present at any given time, which would affect the heterogeneity of unfolded and partially folded protein ensembles.

AUTHOR INFORMATION

Corresponding Author

Jayant B. Udgaonkar − *National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru 560065, India; Indian Institute of Science Education and Research, Pune 411008, India;* orcid.org/0000-0002-7005-224X; Email: jayant@ncbs.res.in, jayant@iiserpune.ac.in

Author

Sandhya Bhatia − *National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru 560065, India; Indian Institute of Science Education and Research, Pune 411008, India;* Present Address: Howard Hughes Medical Institute/Department of Biophysics, UT Southwestern Medical Centre, Dallas 75390, United States; orcid.org/0000-0002-0926-8287

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.chemrev.1c00704

Notes

The authors declare no competing financial interest.

Biographies

Sandhya Bhatia studied at Miranda House College, Delhi, and obtained her B.Sc. degree in chemistry from the University of Delhi. She then pursued an integrated Ph.D. course (M.Sc. + Ph.D.) in Biological sciences (under the supervision of Prof. Jayant B. Udgaonkar) from the National Centre for Biological Sciences, TIFR, which was completed in May 2021. In August 2021, she joined the laboratory of Prof. Michael K. Rosen as a postdoctoral researcher, at Howard Hughes Medical Institute and Department of Biophysics, University of Texas Southwestern Medical Centre, Dallas.

Jayant Udgaonkar studied at St Xavier's College Mumbai and obtained his B.Sc. degree in chemistry from the University of Mumbai. After obtaining his M.Sc. degree in Chemistry from the Indian Institute of Technology, Chennai in 1981, he went on to obtain his Ph.D. in biochemistry (under the supervision of Prof. G. P. Hess) from Cornell University in 1986, with a minor in applied physics. From 1986 to 1989, he was a postdoctoral fellow in the Department of Biochemistry at Stanford University, working with Prof. R. L. Baldwin. He joined the National Centre for Biological Sciences, Tata Institute of Fundamental Research, in 1990. In late 2017, he took leave from the Tata Institute and joined the Indian Institute of Science Education and Research, Pune.

ACKNOWLEDGMENTS

We thank members of our laboratory, as well as G. Krishnamoorthy, Deepak Dhar, 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) Creighton, T. E. Protein Folding. *Biochem. J.* **1990**, *270* (1), 1−16.
(2) Udgaonkar, J. B. Multiple Routes and Structural Heterogeneity in Protein Folding. *Annu. Rev. Biophys.* **2008**, *37* (1), 489−510.
(3) Dill, K. A.; MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science* **2012**, *338* (6110), 1042−1046.
(4) Finkelstein, A. V. 50+ Years of Protein Folding. *Biochemistry (Mosc.)* **2018**, *83*, S3−S18.
(5) Bigman, L. S.; Levy, Y. Proteins: Molecules Defined by Their Trade-Offs. *Curr. Opin. Struct. Biol.* **2020**, *60*, 50−56.
(6) Go, N. The Consistency Principle in Protein Structure and Pathways of Folding. *Adv. Biophys.* **1984**, *18* (C), 149−164.
(7) Malhotra, P.; Udgaonkar, J. B. How Cooperative Are Protein Folding and Unfolding Transitions? *Protein Sci.* **2016**, *25* (11), 1924−1941.
(8) 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.
(9) Sosnick, T. R.; Barrick, D. The Folding of Single Domain Proteins-Have We Reached a Consensus? *Curr. Opin. Struct. Biol.* **2011**, *21* (1), 12−24.
(10) Muñoz, V.; Campos, L. A.; Sadqi, M. Limited Cooperativity in Protein Folding. *Curr. Opin. Struct. Biol.* **2016**, *36*, 58−66.
(11) Brini, E.; Simmerling, C.; Dill, K. Protein Storytelling through Physics. *Science (80-.)* **2020**, *370* (6520). DOI: 10.1126/science.aaz3041
(12) Dobson, C. M. Protein Folding and Misfolding. *Nature* **2003**, *426* (6968), 884−890.
(13) Kumar, S.; Udgaonkar, J. B. Mechanisms of Amyloid Fibril Formation by Proteins. *Curr. Sci.* **2010**, *98* (5), 639−656.
(14) Eisenberg, D.; Jucker, M. The Amyloid State of Proteins in Human Diseases. *Cell* **2012**, *148* (6), 1188−1203.
(15) Ramachandran, G.; Udgaonkar, J. B. Mechanistic Studies Unravel the Complexity Inherent in Tau Aggregation Leading to Alzheimer's Disease and the Tauopathies. *Biochemistry* **2013**, *52* (24), 4107−4126.
(16) Singh, J.; Udgaonkar, J. B. Molecular Mechanism of the Misfolding and Oligomerization of the Prion Protein: Current Understanding and Its Implications. *Biochemistry* **2015**, *54* (29), 4431−4442.
(17) Chiti, F.; Dobson, C. M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu. Rev. Biochem.* **2017**, *86* (1), 27−68.
(18) Koga, N.; Tatsumi-Koga, R.; Liu, G.; Xiao, R.; Acton, T. B.; Montelione, G. T.; Baker, D. Principles for Designing Ideal Protein Structures. *Nature* **2012**, *491* (7423), 222−227.
(19) Compiani, M.; Capriotti, E. Computational and Theoretical Methods for Protein Folding. *Biochemistry* **2013**, *52* (48), 8601−8624.
(20) Udgaonkar, J. B. Polypeptide Chain Collapse and Protein Folding. *Arch. Biochem. Biophys.* **2013**, *531* (1−2), 24−33.
(21) Ptitsyn, O. B. Stages in the Mechanism of Self-Organization of Protein Molecules. *Dokl. Akad. Nauk SSSR* **1973**, *210* (5), 1213−1215.
(22) Kim, P. S.; Baldwin, R. L. Specific Intermediates in the Folding Reactions of Small Proteins and the Mechanism of Protein Folding. *Annu. Rev. Biochem.* **1982**, *51*, 459−489.

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

(23) Udgaonkar, J. B.; Baldwin, R. L. NMR Evidence for an Early Framework Intermediate on the Folding Pathway of Ribonuclease A. *Nature* **1988**, *335* (6192), 694−699.
(24) Karplus, M.; Weaver, D. L. Diffusion-Collision Model for Protein Folding. *Biopolymers* **1979**, *18* (6), 1421−1437.
(25) Karplus, M.; Weaver, D. L. Protein Folding Dynamics: The Diffusion-Collision Model and Experimental Data. *Protein Sci.* **1994**, *3* (4), 650−668.
(26) Bashford, D.; Cohen, F. E.; Karplus, M.; Kuntz, I. D.; Weaver, D. L. Diffusion-Collision Model for the Folding Kinetics of Myoglobin. *Proteins Struct. Funct. Genet.* **1988**, *4* (3), 211−227.
(27) Baldwin, R. L.; Rose, G. D. Is Protein Folding Hierarchic? I. Local Structure and Peptide Folding. *Trends Biochem. Sci.* **1999**, *24* (1), 26−33.
(28) Baldwin, R. L.; Rose, G. D. Is Protein Folding Hierarchic? II. Folding Intermediates and Transition States. *Trends Biochem. Sci.* **1999**, *24* (2), 77−83.
(29) Kuwajima, K.; Yamaya, H.; Miwa, S.; Sugai, S.; Nagamura, T. Rapid Formation of Secondary Structure Framework in Protein Folding Studied by Stopped-Flow Circular Dichroism. *FEBS Lett.* **1987**, *221* (1), 115−118.
(30) Kuwajima, K.; Semisotnov, G. V.; Finkelstein, A. V.; Sugai, S.; Ptitsyn, O. B. Secondary Structure of Globular Proteins at the Early and the Final Stages in Protein Folding. *FEBS Lett.* **1993**, *334* (3), 265−268.
(31) Matthews, C. R. Pathways of protein folding. *Annu. Rev. Biochem.* **1993**, *62* (1), 653−683.
(32) Varley, P.; Gronenborn, A. M.; Christensen, H.; Wingfield, P. T.; Pain, R. H.; Clore, G. M. Kinetics of Folding of the All-β Sheet Protein Interleukin-1β. *Science (80-.).* **1993**, *260* (5111), 1110−1113.
(33) Chamberlain, A. K.; Marqusee, S. Comparison of Equilibrium and Kinetic Approaches for Determining Protein Folding Mechanisms. *Adv. Protein Chem.* **2000**, *53*, 283−328.
(34) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How Fast-Folding Proteins Fold. *Science.* **2011**, *334* (6055), 517−520.
(35) Rollins, G. C.; Dill, K. A. General Mechanism of Two-State Protein Folding Kinetics. *J. Am. Chem. Soc.* **2014**, *136* (32), 11420−11427.
(36) Englander, S. W.; Mayne, L. The Nature of Protein Folding Pathways. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (45), 15873−15880.
(37) Englander, S. W.; Mayne, L. The Case for Defined Protein Folding Pathways. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (31), 8253−8258.
(38) Wetlaufer, D. B. Nucleation, Rapid Folding, and Globular Intrachain Regions in Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1973**, *70* (3), 697−701.
(39) Matheson, R. R.; Scheraga, H. A. A Method for Predicting Nucleation Sites for Protein Folding Based on Hydrophobic Contacts. *Macromolecules* **1978**, *11* (4), 819−829.
(40) Tsong, T. Y.; Baldwin, R. L.; McPhie, P.; Elson, E. L. A Sequential Model of Nucleation-Dependent Protein Folding: Kinetic Studies of Ribonuclease A. *J. Mol. Biol.* **1972**, *63* (3), 453−469.
(41) Itzhaki, L. S.; Otzen, D. E.; Fersht, A. R. The Structure of the Transition State for Folding of Chymotrypsin Inhibitor 2 Analysed by Protein Engineering Methods: Evidence for a Nucleation-Condensation Mechanism for Protein Folding. *J. Mol. Biol.* **1995**, *254* (2), 260−288.
(42) Fersht, A. R. Nucleation Mechanisms in Protein Folding. *Curr. Opin. Struct. Biol.* **1997**, *7* (1), 3−9.
(43) Sánchez, I. E.; Kiefhaber, T. Evidence for Sequential Barriers and Obligatory Intermediates in Apparent Two-State Protein Folding. *J. Mol. Biol.* **2003**, *325* (2), 367−376.
(44) Daggett, V.; Fersht, A. R. Is There a Unifying Mechanism for Protein Folding? *Trends Biochem. Sci.* **2003**, *28* (1), 18−25.
(45) Abkevich, V. I.; Gutin, A. M.; Shakhnovich, E. I. Free Energy Landscape for Protein Folding Kinetics: Intermediates, Traps, and Multiple Pathways in Theory and Lattice Model Simulations. *J. Chem. Phys.* **1994**, *101* (7), 6052−6062.
(46) Guo, Z.; Thirumalai, D. Kinetics of Protein Folding: Nucleation Mechanism, Time Scales, and Pathways. *Biopolymers* **1995**, *36* (1), 83−102.
(47) Faísca, P. F. N. The Nucleation Mechanism of Protein Folding: A Survey of Computer Simulation Studies. *J. Phys.: Condens. Matter* **2009**, *21* (37), 373102.
(48) Thirumalai, D.; O'Brien, E. P.; Morrison, G.; Hyeon, C. Theoretical Perspectives on Protein Folding. *Annu. Rev. Biophys.* **2010**, *39* (1), 159−183.
(49) Fersht, A. R.; Daggett, V. Protein Folding and Unfolding at Atomic Resolution. *Cell* **2002**, *108* (4), 573−582.
(50) Mayor, U.; Guydosh, N. R.; Johnson, C. M.; Grossmann, J. G.; Sato, S.; Jas, G. S.; Freund, S. M. V.; Alonso, D. O. V.; Daggett, V.; Fersht, A. R. The Complete Folding Pathway of a Protein from Nanoseconds to Microseconds. *Nature* **2003**, *421* (6925), 863−867.
(51) Robson, B.; Pain, R. H. Analysis of the Code Relating Sequence to Conformation in Proteins: Possible Implications for the Mechanism of Formation of Helical Regions. *J. Mol. Biol.* **1971**, *58* (1), 237−257.
(52) Dill, K. A. Theory for the Folding and Stability of Globular Proteins. *Biochemistry* **1985**, *24* (6), 1501−1509.
(53) Agashe, V. R.; Shastry, M. C. R.; Udgaonkar, J. B. Initial Hydrophobic Collapse in the Folding of Barstar. *Nature.* **1995**, *377*, 754−757.
(54) Finkelstein, A. V.; Ptitsyn, O. B. Why Do Globular Proteins Fit the Limited Set of Folding Patterns? *Prog. Biophys. Mol. Biol.* **1987**, *50* (3), 171−190.
(55) Dill, K. A. Dominant Forces in Protein Folding. *Biochemistry* **1990**, *29* (31), 7133−7155.
(56) Alonso, D. O. V.; Dill, K. A. Solvent Denaturation and Stabilization of Globular Proteins. *Biochemistry* **1991**, *30* (24), 5974−5985.
(57) Karplus, M.; Shakhnovich, E. In *Protein Folding*; Freeman: New York, 1992; Chapter 4, pp 127−196.
(58) Gutin, A. M.; Abkevich, V. I.; Shakhnovich, E. I. Is Burst Hydrophobic Collapse Necessary for Protein Folding? *Biochemistry* **1995**, *34* (9), 3066−3076.
(59) Flory, P. J. *Principles of Polymer Chemistry*; Cornell University Press: Ithaca, NY, 1953; Vol. *688*.
(60) de Gennes, P. G. Collapse of a Polymer Chain in Poor Solvents. *J. Phys. Lett.* **1975**, *36* (3), 55−57.
(61) Linderstrøm-Lang, K.; Shellman, J. A. Protein Structure and Enzyme Activity. In *The Enzymes*; Academic: New York, 1959; pp 443−510.
(62) Frauenfelder, H.; Parak, F.; Young, R. D. Conformational Substates in Proteins. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 451−479.
(63) Henzler-Wildman, K.; Kern, D. Dynamic Personalities of Proteins. *Nature* **2007**, *450* (7172), 964−972.
(64) Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D. A Unified Model of Protein Dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (13), 5129−5134.
(65) 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.
(66) Sekhar, A.; Kay, L. E. An NMR View of Protein Dynamics in Health and Disease. *Annu. Rev. Biophys.* **2019**, *48* (1), 297−319.
(67) Miller, M. D.; Phillips, G. N. Moving beyond Static Snapshots: Protein Dynamics and the Protein Data Bank. *J. Biol. Chem.* **2021**, *296*, 100749.
(68) Karplus, M.; Kuriyan, J. Molecular Dynamics and Protein Function. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6679−6685.
(69) Naganathan, A. N.; Doshi, U.; Fung, A.; Sadqi, M.; Muñoz, V. Dynamics, Energetics, and Structure in Protein Folding. *Biochemistry* **2006**, *45* (28), 8466−8475.
(70) Kerns, S. J.; Agafonov, R. V.; Cho, Y.-J.; Pontiggia, F.; Otten, R.; Pachov, D. V.; Kutter, S.; Phung, L. A.; Murphy, P. N.; Thai, V.; et al.

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

The Energy Landscape of Adenylate Kinase during Catalysis. *Nat. Struct. Mol. Biol.* **2015**, *22* (2), 124−131.
(71) Narayan, A.; Campos, L. A.; Bhatia, S.; Fushman, D.; Naganathan, A. N. Graded Structural Polymorphism in a Bacterial Thermosensor Protein. *J. Am. Chem. Soc.* **2017**, *139* (2), 792−802.
(72) Alderson, T. R.; Kay, L. E. NMR Spectroscopy Captures the Essential Role of Dynamics in Regulating Biomolecular Function. *Cell* **2021**, *184* (3), 577−595.
(73) Campbell, E.; Kaltenbach, M.; Correy, G. J.; Carr, P. D.; Porebski, B. T.; Livingstone, E. K.; Afriat-Jurnou, L.; Buckle, A. M.; Weik, M.; Hollfelder, F.; et al. The Role of Protein Dynamics in the Evolution of New Enzyme Function. *Nat. Chem. Biol.* **2016**, *12* (11), 944−950.
(74) Campitelli, P.; Modi, T.; Kumar, S.; Ozkan, S. B. The Role of Conformational Dynamics and Allostery in Modulating Protein Evolution. *Annu. Rev. Biophys.* **2020**, *49* (1), 267−288.
(75) Gruebele, M.; Dave, K.; Sukenik, S. Globular Protein Folding In Vitro and In Vivo. *Annu. Rev. Biophys.* **2016**, *45* (1), 233−251.
(76) Anunciado, D. B.; Nyugen, V. P.; Hurst, G. B.; Doktycz, M. J.; Urban, V.; Langan, P.; Mamontov, E.; O'Neill, H. In Vivo Protein Dynamics on the Nanometer Length Scale and Nanosecond Time Scale. *J. Phys. Chem. Lett.* **2017**, *8* (8), 1899−1904.
(77) Boob, M.; Wang, Y.; Gruebele, M. Proteins: 'Boil 'Em, Mash 'Em, Stick 'Em in a Stew.'. *J. Phys. Chem. B* **2019**, *123*, 8341.
(78) Benkovic, S. J.; Hammes, G. G.; Hammes-Schiffer, S. Free-Energy Landscape of Enzyme Catalysis †. *Biochemistry* **2008**, *47* (11), 3317−3321.
(79) Gruebele, M. Protein Dynamics in Simulation and Experiment. *J. Am. Chem. Soc.* **2014**, *136* (48), 16695−16697.
(80) Pierce, M. M.; Nall, B. T. Coupled Kinetic Traps in Cytochrome c Folding: His-Heme Misligation and Proline Isomerization. *J. Mol. Biol.* **2000**, *298* (5), 955−969.
(81) Wani, A. H.; Udgaonkar, J. B. Native State Dynamics Drive the Unfolding of the SH3 Domain of PI3 Kinase at High Denaturant Concentration. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (49), 20711−20716.
(82) Chung, H. S.; Piana-Agostinetti, S.; Shaw, D. E.; Eaton, W. A. Structural Origin of Slow Diffusion in Protein Folding. *Science (80-.).* **2015**, *349* (6255), 1504−1510.
(83) 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.
(84) Goluguri, R. R.; Sen, S.; Udgaonkar, J. B. Microsecond Sub-Domain Motions and the Folding and Misfolding of the Mouse Prion Protein. *eLife* **2019**, *8*, 1−25.
(85) Nath, U.; Udgaonkar, J. B. Folding of Tryptophan Mutants of Barstar: Evidence for an Initial Hydrophobic Collapse on the Folding Pathway. *Biochemistry* **1997**, *36* (28), 8602−8610.
(86) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. Theory of protein folding: The Energy Landscape Perspective. *Annu. Rev. Phys. Chem.* **1997**, *48* (1), 545−600.
(87) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. The Protein Folding Problem. *Annu. Rev. Biophys.* **2008**, *37* (1), 289−316.
(88) England, J. L.; Haran, G. Role of Solvation Effects in Protein Denaturation: From Thermodynamics to Single Molecules and Back. *Annu. Rev. Phys. Chem.* **2011**, *62* (1), 257−277.
(89) Gianni, S.; Jemth, P. Protein Folding: Vexing Debates on a Fundamental Problem. *Biophys. Chem.* **2016**, *212*, 17−21.
(90) Eaton, W. A.; Wolynes, P. G. Theory, Simulations, and Experiments Show That Proteins Fold by Multiple Pathways. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (46), E9759−E9760.
(91) Ferreiro, D. U.; Komives, E. A.; Wolynes, P. G. Frustration, Function and Folding. *Curr. Opin. Struct. Biol.* **2018**, *48*, 68−73.
(92) Schuler, B. Perspective: Chain Dynamics of Unfolded and Intrinsically Disordered Proteins from Nanosecond Fluorescence Correlation Spectroscopy Combined with Single-Molecule FRET. *J. Chem. Phys.* **2018**, *149* (1), No. 010901.
(93) 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.
(94) Ivankov, D. N.; Finkelstein, A. V. Solution of Levinthal's Paradox and a Physical Theory of Protein Folding Times. *Biomolecules* **2020**, *10* (2), 250.
(95) Lapidus, L. J. The Road Less Traveled in Protein Folding: Evidence for Multiple Pathways. *Curr. Opin. Struct. Biol.* **2021**, *66*, 83−88.
(96) Nassar, R.; Dignon, G. L.; Razban, R. M.; Dill, K. A. The Protein Folding Problem: The Role of Theory. *J. Mol. Biol.* **2021**, *433* (20), 167126.
(97) Rose, G. D. Protein Folding - Seeing Is Deceiving. *Protein Sci.* **2021**, *30* (8), 1606−1616.
(98) Haas, E.; Wilchek, M.; Katchalski-Katzir, E.; Steinberg, I. Z. Distribution of End-to-End Distances of Oligopeptides in Solution as Estimated by Energy Transfer. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72* (5), 1807−1811.
(99) Krishnamoorthy, G. Fluorescence Spectroscopy for Revealing Mechanisms in Biology: Strengths and Pitfalls. *J. Biosci.* **2018**, *43* (3), 555−567.
(100) Schuler, B.; Soranno, A.; Hofmann, H.; Nettels, D. Single-Molecule FRET Spectroscopy and the Polymer Physics of Unfolded and Intrinsically Disordered Proteins. *Annu. Rev. Biophys.* **2016**, *45*, 207−231.
(101) Bhatia, S.; Krishnamoorthy, G.; Udgaonkar, J. B. Site-Specific Time-Resolved FRET Reveals Local Variations in the Unfolding Mechanism in an Apparently Two-State Protein Unfolding Transition. *Phys. Chem. Chem. Phys.* **2018**, *20* (5), 3216−3232.
(102) Dill, K. A.; Shortle, D. Denatured States of Proteins. *Annu. Rev. Biochem.* **1991**, *60* (1), 795−825.
(103) Kohn, J. E.; Millett, I. S.; Jacob, J.; Zagrovic, B.; Dillon, T. M.; Cingel, N.; Dothager, R. S.; Seifert, S.; Thiyagarajan, P.; Sosnick, T. R.; et al. Random-Coil Behavior and the Dimensions of Chemically Unfolded Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (34), 12491−12496.
(104) Sherman, E.; Haran, G. Coil-Globule Transition in the Denatured State of a Small Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (31), 11539−11543.
(105) Pitard, E.; Orland, H. Dynamics of the Swelling or Collapse of a Homopolymer. *Europhys. Lett.* **1998**, *41* (4), 467−472.
(106) Halperin, a; Goldbart, P. Early Stages of Homopolymer Collapse. *Phys. Rev. E. Stat. Phys. Plasmas. Fluids. Relat. Interdiscip. Topics* **2000**, *61* (1), 565−573.
(107) Le Guillou, J. C.; Zinn-Justin, J. Critical Exponents for the N-Vector Model in Three Dimensions from Field Theory. *Phys. Rev. Lett.* **1977**, *39* (2), 95−98.
(108) Plaxco, K. W.; Millett, I. S.; Segel, D. J.; Doniach, S.; Baker, D. Chain Collapse Can Occur Concomitantly with the Rate-Limiting Step in Protein Folding. *Nat. Struct. Biol.* **1999**, *6* (6), 554−556.
(109) Kimura, T.; Uzawa, T.; Ishimori, K.; Morishima, I.; Takahashi, S.; Konno, T.; Akiyama, S.; Fujisawa, T. Specific Collapse Followed by Slow Hydrogen-Bond Formation of -Sheet in the Folding of Single-Chain Monellin. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (8), 2748−2753.
(110) Kathuria, S. V.; Kayatekin, C.; Barrea, R.; Kondrashkina, E.; Graceffa, R.; Guo, L.; Nobrega, R. P.; Chakravarthy, S.; Matthews, C. R.; Irving, T. C.; et al. Microsecond Barrier-Limited Chain Collapse Observed by Time-Resolved FRET and SAXS. *J. Mol. Biol.* **2014**, *426* (9), 1980−1994.
(111) 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. Innovative Scattering Analysis Shows That Hydrophobic Disordered Proteins Are Expanded in Water. *Science (80-.).* **2017**, *358* (6360), 238−241.
(112) 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.

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

(113) Aznauryan, M.; Delgado, L.; Soranno, A.; Nettels, D.; Huang, J.; Labhardt, A. M.; Grzesiek, S.; Schuler, B. Comprehensive Structural and Dynamical View of an Unfolded Protein from the Combination of Single-Molecule FRET, NMR, and SAXS. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (37), E5389−E5398.
(114) 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.
(115) Fitzkee, N. C.; Rose, G. D. Reassessing Random-Coil Statistics in Unfolded Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (34), 12497−12502.
(116) 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* (25), 12301−12310.
(117) Bhatia, S.; Krishnamoorthy, G.; Udgaonkar, J. B. Resolving Site-Specific Heterogeneity of the Unfolded State under Folding Conditions. *J. Phys. Chem. Lett.* **2021**, *12* (13), 3295−3302.
(118) Sadqi, M.; Lapidus, L. J.; Munoz, V. How Fast Is Protein Hydrophobic Collapse? *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (21), 12117−12122.
(119) Nettels, D.; Muller-Spath, S.; Kuster, F.; Hofmann, H.; Haenni, D.; Ruegger, S.; Reymond, L.; Hoffmann, A.; Kubelka, J.; Heinz, B.; Gast, K.; Best, R. B.; Schuler, B.; et al. Single-Molecule Spectroscopy of the Temperature-Induced Collapse of Unfolded Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (49), 20740−20745.
(120) Aznauryan, M.; Nettels, D.; Holla, A.; Hofmann, H.; Schuler, B. Single-Molecule Spectroscopy of Cold Denaturation and the Temperature-Induced Collapse of Unfolded Proteins. *J. Am. Chem. Soc.* **2013**, *135* (38), 14040−14043.
(121) Stenzoski, N. E.; Luan, B.; Holehouse, A. S.; Raleigh, D. P. The Unfolded State of the C-Terminal Domain of L9 Expands at Low but Not at Elevated Temperatures. *Biophys. J.* **2018**, *115* (4), 655−663.
(122) Wuttke, R.; Hofmann, H.; Nettels, D.; Borgia, M. B.; Mittal, J.; Best, R. B.; Schuler, B. Temperature-Dependent Solvation Modulates the Dimensions of Disordered Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (14), 5213−5218.
(123) Chandler, D. Interfaces and the Driving Force of Hydrophobic Assembly. *Nature* **2005**, *437* (7059), 640−647.
(124) Reddy, G.; Thirumalai, D. Collapse Precedes Folding in Denaturant-Dependent Assembly of Ubiquitin. *J. Phys. Chem. B* **2017**, *121* (5), 995−1009.
(125) Hicks, A.; Zhou, H.-X. Temperature-Induced Collapse of a Disordered Peptide Observed by Three Sampling Methods in Molecular Dynamics Simulations. *J. Chem. Phys.* **2018**, *149* (7), No. 072313.
(126) Zerze, G. H.; Best, R. B.; Mittal, J. Sequence- and Temperature-Dependent Properties of Unfolded and Disordered Proteins from Atomistic Simulations. *J. Phys. Chem. B* **2015**, *119* (46), 14622−14630.
(127) Ellison, P. A.; Cavagnero, S. Role of Unfolded State Heterogeneity and En-Route Ruggedness in Protein Folding Kinetics. *Protein Sci.* **2006**, *15* (3), 564−582.
(128) Schuler, B.; Lipman, E. A.; Eaton, W. A. Probing the Free-Energy Surface for Protein Folding with Single-Molecule Fluorescence Spectroscopy. *Nature* **2002**, *419* (6908), 743−747.
(129) Kuzmenkina, E. V.; Heyes, C. D.; Nienhaus, G. U. Single-Molecule Förster Resonance Energy Transfer Study of Protein Dynamics under Denaturing Conditions. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (43), 15471−15476.
(130) 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.
(131) Kim, T. W.; Lee, S. J.; Jo, J.; Kim, J. G.; Ki, H.; Kim, C. W.; Cho, K. H.; Choi, J.; Lee, J. H.; Wulff, M.; et al. Protein Folding from Heterogeneous Unfolded State Revealed by Time-Resolved X-Ray Solution Scattering. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (26), 14996−15005.
(132) Soranno, A.; Holla, A.; Dingfelder, F.; Nettels, D.; Makarov, D. E.; Schuler, B. Integrated View of Internal Friction in Unfolded Proteins from Single-Molecule FRET, Contact Quenching, Theory, and Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (10), E1833−E1839.
(133) Hagen, S. J.; Hofrichter, J.; Szabo, A.; Eaton, W. A. Diffusion-Limited Contact Formation in Unfolded Cytochrome c: Estimating the Maximum Rate of Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (21), 11615−11617.
(134) Schwarzinger, S.; Wright, P. E.; Dyson, H. J. Molecular Hinges in Protein Folding: The Urea-Denatured State of Apomyoglobin †. *Biochemistry* **2002**, *41* (42), 12681−12686.
(135) Kuzmenkina, E. V.; Heyes, C. D.; Ulrich Nienhaus, G. Single-Molecule FRET Study of Denaturant Induced Unfolding of RNase H. *J. Mol. Biol.* **2006**, *357* (1), 313−324.
(136) Moglich, A.; Joder, K.; Kiefhaber, T. End-to-End Distance Distributions and Intrachain Diffusion Constants in Unfolded Polypeptide Chains Indicate Intramolecular Hydrogen Bond Formation. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (33), 12394−12399.
(137) Hoffmann, A.; Kane, A.; Nettels, D.; Hertzog, D. E.; Baumgartel, P.; Lengefeld, J.; Reichardt, G.; Horsley, D. A.; Seckler, R.; Bakajin, O.; et al. Mapping Protein Collapse with Single-Molecule Fluorescence and Kinetic Synchrotron Radiation Circular Dichroism Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (1), 105−110.
(138) Soranno, A.; Longhi, R.; Bellini, T.; Buscaglia, M. Kinetics of Contact Formation and End-to-End Distance Distributions of Swollen Disordered Peptides. *Biophys. J.* **2009**, *96* (4), 1515−1528.
(139) 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.
(140) Borgia, A.; Wensley, B. G.; Soranno, A.; Nettels, D.; Borgia, M. B.; Hoffmann, A.; Pfeil, S. H.; Lipman, E. A.; Clarke, J.; Schuler, B. Localizing Internal Friction along the Reaction Coordinate of Protein Folding by Combining Ensemble and Single-Molecule Fluorescence Spectroscopy. *Nat. Commun.* **2012**, *3*, 1−9.
(141) Chattopadhyay, K.; Elson, E. L.; Frieden, C. The Kinetics of Conformational Fluctuations in an Unfolded Protein Measured by Fluorescence Methods. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (7), 2385−2389.
(142) Chen, H.; Rhoades, E.; Butler, J. S.; Loh, S. N.; Webb, W. W. Dynamics of Equilibrium Structural Fluctuations of Apomyoglobin Measured by Fluorescence Correlation Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (25), 10459−10464.
(143) Voelz, V. A.; Singh, V. R.; Wedemeyer, W. J.; Lapidus, L. J.; Pande, V. S. Unfolded-State Dynamics and Structure of Protein L Characterized by Simulation and Experiment. *J. Am. Chem. Soc.* **2010**, *132* (13), 4702−4709.
(144) Voelz, V. A.; Jäger, M.; Yao, S.; Chen, Y.; Zhu, L.; Waldauer, S. A.; Bowman, G. R.; Friedrichs, M.; Bakajin, O.; Lapidus, L. J.; et al. Slow Unfolded-State Structuring in Acyl-CoA Binding Protein Folding Revealed by Simulation and Experiment. *J. Am. Chem. Soc.* **2012**, *134* (30), 12565−12577.
(145) Gopich, I. V.; Szabo, A. FRET Efficiency Distributions of Multistate Single Molecules. *J. Phys. Chem. B* **2010**, *114* (46), 15221−15226.
(146) Chung, H. S.; Gopich, I. V.; McHale, K.; Cellmer, T.; Louis, J. M.; Eaton, W. A. Extracting Rate Coefficients from Single-Molecule Photon Trajectories and FRET Efficiency Histograms for a Fast-Folding Protein. *J. Phys. Chem. A* **2011**, *115* (16), 3642−3656.
(147) Hoffmann, A.; Nettels, D.; Clark, J.; Borgia, A.; Radford, S. E.; Clarke, J.; Schuler, B. Quantifying Heterogeneity and Conformational Dynamics from Single Molecule FRET of Diffusing Molecules: Recurrence Analysis of Single Particles (RASP). *Phys. Chem. Chem. Phys.* **2011**, *13* (5), 1857.
(148) Merchant, K. a; Best, R. B.; Louis, J. M.; Gopich, I. V.; Eaton, W. a. Characterizing the Unfolded States of Proteins Using Single-

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

Molecule FRET Spectroscopy and Molecular Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (5), 1528−1533.
(149) Saito, M.; Kamonprasertsuk, S.; Suzuki, S.; Nanatani, K.; Oikawa, H.; Kushiro, K.; Takai, M.; Chen, P. T.; Chen, E. H. L.; Chen, R. P. Y.; et al. Significant Heterogeneity and Slow Dynamics of the Unfolded Ubiquitin Detected by the Line Confocal Method of Single-Molecule Fluorescence Spectroscopy. *J. Phys. Chem. B* **2016**, *120* (34), 8818−8829.
(150) Oikawa, H.; Kamagata, K.; Arai, M.; Takahashi, S. Complexity of the Folding Transition of the b Domain of Protein a Revealed by the High-Speed Tracking of Single-Molecule Fluorescence Time Series. *J. Phys. Chem. B* **2015**, *119* (20), 6081−6091.
(151) Deniz, A. A.; Laurence, T. A.; Dahan, M.; Chemla, D. S.; Schultz, P. G.; Weiss, S. Ratiometric Single-Molecule Studies of Freely Diffusing Biomolecules. *Annu. Rev. Phys. Chem.* **2001**, *52* (1), 233−253.
(152) Brandts, J. F.; Halvorson, H. R.; Brennan, M. Consideration of the Possibility That the Slow Step in Protein Denaturation Reactions Is Due to Cis-Trans Isomerism of Proline Residues. *Biochemistry* **1975**, *14* (22), 4953−4963.
(153) Shastry, M. C. R.; Agashe, V. R.; Udgaonkar, J. B. Quantitative Analysis of the Kinetics of Denaturation and Renaturation of Barstar in the Folding Transition Zone. *Protein Sci.* **1994**, *3* (9), 1409−1417.
(154) Schönbrunner, N.; Koller, K. P.; Kiefhaber, T. Folding of the Disulfide-Bonded β-Sheet Protein Tendamistat: Rapid Two-State Folding without Hydrophobic Collapse. *J. Mol. Biol.* **1997**, *268* (2), 526−538.
(155) Wedemeyer, W. J.; Welker, E.; Scheraga, H. A. Proline Cis-Trans Isomerization and Protein Folding. *Biochemistry* **2002**, *41* (50), 14637−14644.
(156) Wu, Y.; Kondrashkina, E.; Kayatekin, C.; Matthews, C. R.; Bilsel, O. Microsecond Acquisition of Heterogeneous Structure in the Folding of a TIM Barrel Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (36), 13367−13372.
(157) Waudby, C. A.; Wlodarski, T.; Karyadi, M. E.; Cassaignau, A. M. E.; Chan, S. H. S.; Wentink, A. S.; Schmidt-Engler, J. M.; Camilloni, C.; Vendruscolo, M.; Cabrita, L. D.; et al. Systematic Mapping of Free Energy Landscapes of a Growing Filamin Domain during Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (39), 9744−9749.
(158) Schmid, F. X.; Baldwin, R. L. The Rate of Interconversion between the Two Unfolded Forms of Ribonuclease A Does Not Depend on Guanidinium Chloride Concentration. *J. Mol. Biol.* **1979**, *133* (2), 285−287.
(159) Osváth, S.; Gruebele, M. Proline Can Have Opposite Effects on Fast and Slow Protein Folding Phases. *Biophys. J.* **2003**, *85* (2), 1215−1222.
(160) Schindler, T.; Mayr, L. M.; Landt, O.; Hahn, U.; Schmid, F. X. The Role of a Trans-Proline in the Folding Mechanism of Ribonuclease T1. *Eur. J. Biochem.* **1996**, *241* (2), 516−524.
(161) Mateos, B.; Conrad-billroth, C.; Schiavina, M.; Beier, A.; Kontaxis, G.; Konrat, R.; Felli, I. C.; Pierattelli, R. The Ambivalent Role of Proline Residues in an Intrinsically Disordered Protein: From Disorder Promoters to Compaction Facilitators. *J. Mol. Biol.* **2020**, *432*, 3093−3111.
(162) 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.
(163) Moors, S. L. C.; Hellings, M.; De Maeyer, M.; Engelborghs, Y.; Ceulemans, A. Tryptophan Rotamers as Evidenced by X-Ray, Fluorescence Lifetimes, and Molecular Dynamics Modeling. *Biophys. J.* **2006**, *91* (3), 816−823.
(164) Darby, N.; Creighton, T. E. Disulfide Bonds in Protein Folding and Stability. In *Protein Stability and Folding*; Humana Press: NJ, 1995; pp 219−252.
(165) Weissman, J. S. All Roads Lead to Rome? The Multiple Pathways of Protein Folding. *Chem. Biol.* **1995**, *2* (5), 255−260.
(166) Welker, E.; Wedemeyer, W. J.; Narayan, M.; Scheraga, H. A. Coupling of Conformational Folding and Disulfide-Bond Reactions in Oxidative Folding of Proteins. *Biochemistry* **2001**, *40* (31), 9059−9064.
(167) Goldenberg, D. P.; Creighton, T. E. Folding Pathway of a Circular Form of Bovine Pancreatic Trypsin Inhibitor. *J. Mol. Biol.* **1984**, *179* (3), 527−545.
(168) Weissman, J.; Kim, P. Reexamination of the Folding of BPTI: Predominance of Native Intermediates. *Science.* **1991**, *253* (5026), 1386−1393.
(169) Weissman, J. S.; Kimt, P. S. Efficient Catalysis of Disulphide Bond Rearrangements by Protein Disulphide Isomerase. *Nature* **1993**, *365* (6442), 185−188.
(170) Kawagoe, S.; Nakagawa, H.; Kumeta, H.; Ishimori, K.; Saio, T. Structural Insight into Proline Cis/Trans Isomerization of Unfolded Proteins Catalyzed by the Trigger Factor Chaperone. *J. Biol. Chem.* **2018**, *293* (39), 15095−15106.
(171) Bhuyan, A. K.; Udgaonkar, J. B. Multiple Kinetic Intermediates Accumulate during the Unfolding of Horse Cytochrome c in the Oxidized State. *Biochemistry* **1998**, *37* (25), 9147−9155.
(172) Sosnick, T. R.; Mayne, L.; Hiller, R.; Englander, S. W. The Barriers in Protein Folding. *Nat. Struct. Biol.* **1994**, *1* (3), 149−156.
(173) Yang, W. Y.; Gruebele, M. Detection-Dependent Kinetics as a Probe of Folding Landscape Microstructure. *J. Am. Chem. Soc.* **2004**, *126* (25), 7758−7759.
(174) Patra, A. K.; Udgaonkar, J. B. Characterization of the Folding and Unfolding Reactions of Single-Chain Monellin: Evidence for Multiple Intermediates and Competing Pathways. *Biochemistry* **2007**, *46* (42), 11727−11743.
(175) 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.
(176) Nöppert, A.; Gast, K.; Müller-Frohne, M.; Zirwer, D.; Damaschun, G. Reduced-Denatured Ribonuclease A Is Not in a Compact State. *FEBS Lett.* **1996**, *380* (1−2), 179−182.
(177) 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.
(178) Sridevi, K.; Udgaonkar, J. B. Surface Expansion Is Independent of and Occurs Faster than Core Solvation during the Unfolding of Barstar. *Biochemistry* **2003**, *42* (6), 1551−1563.
(179) Sinha, K. K.; Udgaonkar, J. B. Dependence of the Size of the Initially Collapsed Form during the Refolding of Barstar on Denaturant Concentration: Evidence for a Continuous Transition. *J. Mol. Biol.* **2005**, *353* (3), 704−718.
(180) Sinha, K. K.; Udgaonkar, J. B. Dissecting the Non-Specific and Specific Components of the Initial Folding Reaction of Barstar by Multi-Site FRET Measurements. *J. Mol. Biol.* **2007**, *370* (2), 385−405.
(181) Schuler, B.; Eaton, W. A. Protein Folding Studied by Single-Molecule FRET. *Curr. Opin. Struct. Biol.* **2008**, *18* (1), 16−26.
(182) 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.
(183) Ferreon, A. C. M.; Deniz, A. A. α-Synuclein Multistate Folding Thermodynamics: Implications for Protein Misfolding and Aggregation. *Biochemistry* **2007**, *46* (15), 4499−4509.
(184) Mukhopadhyay, S.; Krishnan, R.; Lemke, E. A.; Lindquist, S.; Deniz, A. A. A Natively Unfolded Yeast Prion Monomer Adopts an Ensemble of Collapsed and Rapidly Fluctuating Structures. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (8), 2649−2654.
(185) Ferreon, A. C. M.; Moran, C. R.; Gambin, Y.; Deniz, A. A. Single-Molecule Fluorescence Studies of Intrinsically Disordered Proteins. In *Single Molecule Tools: Fluorescence Based Approaches, Part A*; Elsevier Inc., 2010; Vol. *472*, pp 179−204.
(186) Schellman, J. A. Solvent Denaturation. *Biopolymers* **1978**, *17* (5), 1305−1322.
(187) Makhatadze, G. I.; Privalov, P. L. Protein Interactions with Urea and Guanidinium Chloride. *J. Mol. Biol.* **1992**, *226* (2), 491−505.

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

(188) Holehouse, A. S.; Pappu, R. V. Collapse Transitions of Proteins and the Interplay among Backbone, Sidechain, and Solvent Interactions. *Annual Review of Biophysics.* **2018**, *47*, 19−39.
(189) Hamadani, K. M.; Weiss, S. Nonequilibrium Single Molecule Protein Folding in a Coaxial Mixer. *Biophys. J.* **2008**, *95* (1), 352−365.
(190) Buckler, D. R.; Haas, E.; Scheraga, H. A. Analysis of the Structure of Ribonuclease A in Native and Partially Denatured States by Time-Resolved Nonradiative Dynamic Excitation Energy Transfer between Site-Specific Extrinsic Probes. *Biochemistry* **1995**, *34* (49), 15965−15978.
(191) Otosu, T.; Ishii, K.; Oikawa, H.; Arai, M.; Takahashi, S.; Tahara, T. Highly Heterogeneous Nature of the Native and Unfolded States of the B Domain of Protein A Revealed by Two-Dimensional Fluorescence Lifetime Correlation Spectroscopy. *J. Phys. Chem. B* **2017**, *121* (22), 5463−5473.
(192) Bowler, B. E. Residual Structure in Unfolded Proteins. *Curr. Opin. Struct. Biol.* **2012**, *22* (1), 4−13.
(193) Neri, D.; Billeter, M.; Wider, G.; Wuthrich, K. NMR Determination of Residual Structure in a Urea-Denatured Protein, the 434-Repressor. *Science (80-.).* **1992**, *257* (5076), 1559−1563.
(194) Schwalbe, H.; Fiebig, K. M.; Buck, M.; Jones, J. A.; Grimshaw, S. B.; Spencer, A.; Glaser, S. J.; Smith, L. J.; Dobson, C. M. Structural and Dynamical Properties of a Denatured Protein. Heteronuclear 3D NMR Experiments and Theoretical Simulations of Lysozyme in 8 M Urea †. *Biochemistry* **1997**, *36* (29), 8977−8991.
(195) McCarney, E. R.; Werner, J. H.; Bernstein, S. L.; Ruczinski, I.; Makarov, D. E.; Goodwin, P. M.; Plaxco, K. W. Site-Specific Dimensions across a Highly Denatured Protein; a Single Molecule Study. *J. Mol. Biol.* **2005**, *352* (3), 672−682.
(196) Wirmer, J.; Berk, H.; Ugolini, R.; Redfield, C.; Schwalbe, H. Characterization of the Unfolded State of Bovine α-Lactalbumin and Comparison with Unfolded States of Homologous Proteins. *Protein Sci.* **2006**, *15* (6), 1397−1407.
(197) Meier, S.; Strohmeier, M.; Blackledge, M.; Grzesiek, S. Direct Observation of Dipolar Couplings and Hydrogen Bonds across a β-Hairpin in 8 M Urea. *J. Am. Chem. Soc.* **2007**, *129* (4), 754−755.
(198) Huang, J. R.; Grzesiek, S. Ensemble Calculations of Unstructured Proteins Constrained by RDC and PRE Data: A Case Study of Urea-Denatured Ubiquitin. *J. Am. Chem. Soc.* **2010**, *132* (2), 694−705.
(199) Jensen, M. R.; Zweckstetter, M.; Huang, J. R.; Blackledge, M. Exploring Free-Energy Landscapes of Intrinsically Disordered Proteins at Atomic Resolution Using NMR Spectroscopy. *Chem. Rev.* **2014**, *114* (13), 6632−6660.
(200) Pradeep, L.; Udgaonkar, J. B. Effect of Salt on the Urea-Unfolded Form of Barstar Probed by m Value Measurements. *Biochemistry* **2004**, *43* (36), 11393−11402.
(201) Bhavesh, N. S.; Juneja, J.; Udgaonkar, J. B.; Hosur, R. V. Native and Nonnative Conformational Preferences in the Urea-Unfolded State of Barstar. *Protein Sci.* **2004**, *13* (12), 3085−3091.
(202) Sakurai, K.; Yagi, M.; Konuma, T.; Takahashi, S.; Nishimura, C.; Goto, Y. Non-Native α-Helices in the Initial Folding Intermediate Facilitate the Ordered Assembly of the β-Barrel in β-Lactoglobulin. *Biochemistry* **2017**, *56* (36), 4799−4807.
(203) Halloran, K. T.; Wang, Y.; Arora, K.; Chakravarthy, S.; Irving, T. C.; Bilsel, O.; Brooks, C. L.; Matthews, C. R. Frustration and Folding of a TIM Barrel Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (33), 16378−16383.
(204) Shortle, D.; Ackerman, M. S. Persistence of Native-like Topology in a Denatured Protein in 8 M Urea. *Science.* **2001**, *293* (5529), 487−489.
(205) Shortle, D. The Expanded Denatured State: An Ensemble of Conformations Trapped in a Locally Encoded Topological Space. *Adv. Protein Chem.* **2002**, *62*, 1−23.
(206) Klein-Seetharaman, J.; Oikawa, M.; Grimshaw, S. B.; Wirmer, J.; Duchardt, E.; Ueda, T.; Imoto, T.; Smith, L. J.; Dobson, C. M.; Schwalbe, H. Long-Range Interactions within a Nonnative Protein. *Science.* **2002**, *295* (5560), 1719−1722.
(207) Deniz, A. A.; Dahan, M.; Grunwell, J. R.; Ha, T.; Faulhaber, A. E.; Chemla, D. S.; Weiss, S.; Schultz, P. G. Single-Pair Fluorescence Resonance Energy Transfer on Freely Diffusing Molecules: Observation of Förster Distance Dependence and Subpopulations. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (7), 3670−3675.
(208) Chen, H.; Ahsan, S. S.; Santiago-Berrios, M. B.; Abruña, H. D.; Webb, W. W. Mechanisms of Quenching of Alexa Fluorophores by Natural Amino Acids. *J. Am. Chem. Soc.* **2010**, *132* (21), 7244−7245.
(209) Hoefling, M.; Lima, N.; Haenni, D.; Seidel, C. A. M.; Schuler, B.; Grubmüller, H. Structural Heterogeneity and Quantitative FRET Efficiency Distributions of Polyprolines through a Hybrid Atomistic Simulation and Monte Carlo Approach. *PLoS One* **2011**, *6* (5), e19791.
(210) Holden, S. J.; Uphoff, S.; Hohlbein, J.; Yadin, D.; Le Reste, L.; Britton, O. J.; Kapanidis, A. N. Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophys. J.* **2010**, *99* (9), 3102−3111.
(211) Juneja, J.; Udgaonkar, J. B. NMR Studies of Protein Folding. *Curr. Sci.* **2003**, *84* (2), 157−172.
(212) Meng, W.; Luan, B.; Lyle, N.; Pappu, R. V.; Raleigh, D. P. The Denatured State Ensemble Contains Significant Local and Long-Range Structure under Native Conditions: Analysis of the N-Terminal Domain of Ribosomal Protein L9. *Biochemistry* **2013**, *52* (15), 2662−2671.
(213) Ozenne, V.; Noel, J. K.; Heidarsson, P. O.; Brander, S.; Poulsen, F. M.; Jensen, M. R.; Kragelund, B. B.; Blackledge, M.; Danielsson, J. Exploring the Minimally Frustrated Energy Landscape of Unfolded ACBP. *J. Mol. Biol.* **2014**, *426* (3), 722−734.
(214) Yagi-Utsumi, M.; Chandak, M. S.; Yanaka, S.; Hiranyakorn, M.; Nakamura, T.; Kato, K.; Kuwajima, K. Residual Structure of Unfolded Ubiquitin as Revealed by Hydrogen/Deuterium-Exchange 2D NMR. *Biophys. J.* **2020**, *119* (10), 2029−2038.
(215) Chan, H. S.; Dill, K. A. Origins of Structure in Globular Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87* (16), 6388−6392.
(216) 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.
(217) Han, X.; Sit, A.; Christoffer, C.; Chen, S.; Kihara, D. A Global Map of the Protein Shape Universe. *PLOS Comput. Biol.* **2019**, *15* (4), e1006969.
(218) Baldwin, R. L. Energetics of Protein Folding. *J. Mol. Biol.* **2007**, *371* (2), 283−301.
(219) Hvidt, A.; Linderstrøm-Lang, K. Exchange of Hydrogen Atoms in Insulin with Deuterium Atoms in Aqueous Solutions. *Biochim. Biophys. Acta* **1954**, *14*, 574−575.
(220) Saxena, A. M.; Udgaonkar, J. B.; Krishnamoorthy, G. Protein Dynamics Control Proton Transfer from Bulk Solvent to Protein Interior: A Case Study with a Green Fluorescent Protein. *Protein Sci.* **2005**, *14*, 1787−1799.
(221) Ries, J.; Schwarze, S.; Johnson, C. M.; Neuweiler, H. Microsecond Folding and Domain Motions of a Spider Silk Protein Structural Switch. *J. Am. Chem. Soc.* **2014**, *136* (49), 17136−17144.
(222) Englander, S. W.; Mayne, L.; Kan, Z.-Y.; Hu, W. Protein Folding-How and Why: By Hydrogen Exchange, Fragment Separation, and Mass Spectrometry. *Annu. Rev. Biophys.* **2016**, *45*, 135−152.
(223) Privalov, P. L. Stability of Proteins Small Globular Proteins. In *Advances in Protein Chemistry*; Elsevier, 1979; pp 167−241.
(224) Bowman, G. R.; Geissler, P. L. Extensive Conformational Heterogeneity within Protein Cores. *J. Phys. Chem. B* **2014**, *118* (24), 6417−6423.
(225) 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.
(226) Sridevi, K.; Lakshmikanth, G. S.; Krishnamoorthy, G.; Udgaonkar, J. B. Increasing Stability Reduces Conformational Heterogeneity in a Protein Folding Intermediate Ensemble. *J. Mol. Biol.* **2004**, *337* (3), 699−711.

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

(227) Mazouchi, A.; Zhang, Z.; Bahram, A.; Gomes, G. N.; Lin, H.; Song, J.; Chan, H. S.; Forman-Kay, J. D.; Gradinaru, C. C. Conformations of a Metastable SH3 Domain Characterized by SmFRET and an Excluded-Volume Polymer Model. *Biophys. J.* **2016**, *110* (7), 1510−1522.
(228) Bai, Y.; Sosnick, T. R.; Mayne, L.; Englander, S. W. Protein Folding Intermediates: Native-State Hydrogen Exchange. *Science.* **1995**, *269* (5221), 192−197.
(229) Malhotra, P.; Udgaonkar, J. B. High-Energy Intermediates in Protein Unfolding Characterized by Thiol Labeling under Nativelike Conditions. *Biochemistry* **2014**, *53* (22), 3608−3620.
(230) Malhotra, P.; Udgaonkar, J. B. Secondary Structural Change Can Occur Diffusely and Not Modularly during Protein Folding and Unfolding Reactions. *J. Am. Chem. Soc.* **2016**, *138* (18), 5866−5878.
(231) Juneja, J.; Udgaonkar, J. B. Characterization of the Unfolding of Ribonuclease A by a Pulsed Hydrogen Exchange Study: Evidence for Competing Pathways for Unfolding. *Biochemistry* **2002**, *41* (8), 2641−2654.
(232) Doose, S.; Neuweiler, H.; Sauer, M. Fluorescence Quenching by Photoinduced Electron Transfer: A Reporter for Conformational Dynamics of Macromolecules. *ChemPhysChem* **2009**, *10* (9−10), 1389−1398.
(233) Schmidpeter, P. A. M.; Koch, J. R.; Schmid, F. X. Control of Protein Function by Prolyl Isomerization. *Biochim. Biophys. Acta - Gen. Subj.* **2015**, *1850* (10), 1973−1982.
(234) Brichkina, A.; Nguyen, N. T.; Baskar, R.; Wee, S.; Gunaratne, J.; Robinson, R. C.; Bulavin, D. V. Proline Isomerisation as a Novel Regulatory Mechanism for P38MAPK Activation and Functions. *Cell Death Differ.* **2016**, *23* (10), 1592−1601.
(235) Chiti, F.; De Lorenzi, E.; Grossi, S.; Mangione, P.; Giorgetti, S.; Caccialanza, G.; Dobson, C. M.; Merlini, G.; Ramponi, G.; Bellotti, V. A Partially Structured Species of B2-Microglobulin Is Significantly Populated under Physiological Conditions and Involved in Fibrillogenesis. *J. Biol. Chem.* **2001**, *276* (50), 46714−46721.
(236) Mukaiyama, A.; Nakamura, T.; Makabe, K.; Maki, K.; Goto, Y.; Kuwajima, K. Native-State Heterogeneity of B2-Microglobulin as Revealed by Kinetic Folding and Real-Time NMR Experiments. *J. Mol. Biol.* **2013**, *425* (2), 257−272.
(237) Evans, P. A.; Dobson, C. M.; Kautz, R. A.; Hatfull, G.; Fox, R. O. Proline Isomerism in Staphylococcal Nuclease Characterized by NMR and Site-Directed Mutagenesis. *Nature* **1987**, *329* (6136), 266−268.
(238) Zosel, F.; Mercadante, D.; Nettels, D.; Schuler, B. A Proline Switch Explains Kinetic Heterogeneity in a Coupled Folding and Binding Reaction. *Nat. Commun.* **2018**, *9* (1). DOI: 10.1038/s41467-018-05725-0
(239) Li, H.; Frieden, C. NMR Studies of 4−19F-Phenylalanine-Labeled Intestinal Fatty Acid Binding Protein: Evidence for Conformational Heterogeneity in the Native State. *Biochemistry* **2005**, *44* (7), 2369−2377.
(240) Serrano, A. L.; Bilsel, O.; Gai, F. Native State Conformational Heterogeneity of HP35 Revealed by Time-Resolved FRET. *J. Phys. Chem. B* **2012**, *116* (35), 10631−10638.
(241) Aviram, H. Y.; Pirchi, M.; Mazal, H.; Barak, Y.; Riven, I.; Haran, G. Direct Observation of Ultrafast Large-Scale Dynamics of an Enzyme under Turnover Conditions. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (13), 3243−3248.
(242) Grossman-Haham, I.; Rosenblum, G.; Namani, T.; Hofmann, H. Slow Domain Reconfiguration Causes Power-Law Kinetics in a Two-State Enzyme. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (3), 513−518.
(243) Lee, J. H.; Zhang, D.; Hughes, C.; Okuno, Y.; Sekhar, A.; Cavagnero, S. Heterogeneous Binding of the SH3 Client Protein to the DnaK Molecular Chaperone. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (31), E4206−E4215.
(244) Noguera, M. E.; Vazquez, D. S.; Ferrer-Sueta, G.; Agudelo, W. A.; Howard, E.; Rasia, R. M.; Manta, B.; Cousido-Siah, A.; Mitschler, A.; Podjarny, A.; Santos, J. Structural Variability of E. Coli Thioredoxin Captured in the Crystal Structures of Single-Point Mutants. *Sci. Rep.* **2017**, *7*, 42343.
(245) Thirumalai, D. From Minimal Models to Real Proteins: Time Scales for Protein Folding Kinetics. *J. Phys. I* **1995**, *5* (11), 1457−1467.
(246) Welker, E.; Maki, K.; Shastry, M. C. R.; Juminaga, D.; Bhat, R.; Scheraga, H. A.; Roder, H. 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.* **2004**, *101* (51), 17681−17686.
(247) 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.
(248) Graziano, G. Is Water a Good Solvent for the Denatured State of Globular Proteins? *Chem. Phys. Lett.* **2020**, *759*, 137949.
(249) Maity, H.; Reddy, G. Folding of Protein L with Implications for Collapse in the Denatured State Ensemble. *J. Am. Chem. Soc.* **2016**, *138* (8), 2609−2616.
(250) 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.
(251) 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.* **2018**, *430* (4), 465−478.
(252) Best, R. B. Emerging Consensus on the Collapse of Unfolded and Intrinsically Disordered Proteins in Water. *Curr. Opin. Struct. Biol.* **2020**, *60*, 27−38.
(253) 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.
(254) 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.
(255) Millet, I. S.; Townsley, L. E.; Chiti, F.; Doniach, S.; Plaxco, K. W. Equilibrium Collapse and the Kinetic "foldability" of Proteins. *Biochemistry* **2002**, *41* (1), 321−325.
(256) Magg, C.; Schmid, F. X. Rapid Collapse Precedes the Fast Two-State Folding of the Cold Shock Protein. *J. Mol. Biol.* **2004**, *335* (5), 1309−1323.
(257) 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.
(258) Lapidus, L. J.; Yao, S.; McGarrity, K. S.; Hertzog, D. E.; Tubman, E.; Bakajin, O. Protein Hydrophobic Collapse and Early Folding Steps Observed in a Microfluidic Mixer. *Biophys. J.* **2007**, *93* (1), 218−224.
(259) 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.
(260) Shastry, M. C. R.; Roder, H. Evidence for Barrier-Limited Protein Folding Kinetics on the Microsecond Time Scale. *Nat. Struct. Biol.* **1998**, *5* (5), 385−392.
(261) Rosen, L. E.; Kathuria, S. V.; Matthews, C. R.; Bilsel, O.; Marqusee, S. Non-Native Structure Appears in Microseconds during the Folding of E. Coli RNase H. *J. Mol. Biol.* **2015**, *427* (2), 443−453.
(262) Ansari, A.; Jones, C.; Henry, E.; Hofrichter, J.; Eaton, W. The Role of Solvent Viscosity in the Dynamics of Protein Conformational Changes. *Science (80-.).* **1992**, *256* (5065), 1796−1798.
(263) Pabit, S. A.; Roder, H.; Hagen, S. J. Internal Friction Controls the Speed of Protein Folding from a Compact Configuration. *Biochemistry* **2004**, *43* (39), 12532−12538.
(264) Nettels, D.; Gopich, I. V.; Hoffmann, A.; Schuler, B. Ultrafast Dynamics of Protein Collapse from Single-Molecule Photon Statistics. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (8), 2655−2660.
(265) Jacob, J.; Krantz, B.; Dothager, R. S.; Thiyagarajan, P.; Sosnick, T. R. Early Collapse Is Not an Obligate Step in Protein Folding. *J. Mol. Biol.* **2004**, *338* (2), 369−382.

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

(266) 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.
(267) 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.
(268) Arai, M.; Iwakura, M.; Matthews, C. R.; Bilsel, O. Microsecond Subdomain Folding in Dihydrofolate Reductase. *J. Mol. Biol.* **2011**, *410* (2), 329−342.
(269) Ziv, G.; Thirumalai, D.; Haran, G. Collapse Transition in Proteins. *Phys. Chem. Chem. Phys.* **2009**, *11* (1), 83−93.
(270) 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.
(271) Riback, J. A.; Bowman, M. A.; Zmyslowski, A. M.; Plaxco, K. W.; Clark, P. L.; Sosnick, T. R. Commonly Used FRET Fluorophores Promote Collapse of an Otherwise Disordered Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (18), 8889−8894.
(272) 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.
(273) Fuertes, G.; Banterle, N.; Ruff, K. M.; Chowdhury, A.; Mercadante, D.; Koehler, C.; Kachala, M.; Estrada Girona, G.; Milles, S.; Mishra, A.; Onck, P. R.; Grater, F.; Esteban-Martin, S.; Pappu, R. V.; Svergun, D. I.; Lemke, E. A. Decoupling of Size and Shape Fluctuations in Heteropolymeric Sequences Reconciles Discrepancies in SAXS vs. FRET Measurements. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (31), E6342−E6351.
(274) Ruff, K. M.; Holehouse, A. S. SAXS versus FRET: A Matter of Heterogeneity? *Biophys. J.* **2017**, *113* (5), 971−973.
(275) 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.
(276) Fuertes, G.; Banterle, N.; Ruff, K. M.; Chowdhury, A.; Pappu, R. V.; Svergun, D. I.; Lemke, E. A. Comment on "Innovative Scattering Analysis Shows That Hydrophobic Disordered Proteins Are Expanded in Water. *Science* **2018**, *361* (6405), eaau8230.
(277) Moses, D.; Yu, F.; Ginell, G. M.; Shamoon, N. M.; Koenig, P. S.; Holehouse, A. S.; Sukenik, S. Revealing the Hidden Sensitivity of Intrinsically Disordered Proteins to Their Chemical Environment. *J. Phys. Chem. Lett.* **2020**, *11*, 10131−10136.
(278) Yamada, S.; Bouley Ford, N. D.; Keller, G. E.; Ford, W. C.; Gray, H. B.; Winkler, J. R. Snapshots of a Protein Folding Intermediate. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (5), 1606−1610.
(279) Navon, A.; Ittah, V.; Landsman, P.; Scheraga, H. A.; Haas, E. Distributions of Intramolecular Distances in the Reduced and Denatured States of Bovine Pancreatic Ribonuclease A. Folding Initiation Structures in the C-Terminal Portions of the Reduced Protein. *Biochemistry* **2001**, *40* (1), 105−118.
(280) Flanagan, J. M.; Kataoka, M.; Shortle, D.; Engelman, D. M. Truncated Staphylococcal Nuclease Is Compact but Disordered. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (2), 748−752.
(281) Lietzow, M. A.; Jamin, M.; Jane Dyson, H.; Wright, P. E. Mapping Long-Range Contacts in a Highly Unfolded Protein. *J. Mol. Biol.* **2002**, *322* (4), 655−662.
(282) Ptitsyn, O. B.; Uversky, V. N. The Molten Globule Is a Third Thermodynamical State of Protein Molecules. *FEBS Lett.* **1994**, *341* (1), 15−18.
(283) 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.
(284) 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.
(285) Mizukami, T.; Xu, M.; Cheng, H.; Roder, H.; Maki, K. Nonuniform Chain Collapse during Early Stages of Staphylococcal Nuclease Folding Detected by Fluorescence Resonance Energy Transfer and Ultrarapid Mixing Methods. *Protein Sci.* **2013**, *22* (10), 1336−1348.
(286) 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 -Helical Content and Compactness. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (5), 1171−1176.
(287) 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.
(288) Morar, A. S.; Olteanu, A.; Young, G. B.; Pielak, G. J. Solvent-Induced Collapse of Alpha-Synuclein and Acid-Denatured Cytochrome C. *Protein Sci.* **2001**, *10* (11), 2195−2199.
(289) Camacho, C. J.; Thirumalai, D. Kinetics and Thermodynamics of Folding in Model Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (13), 6369−6372.
(290) Religa, T. L.; Markson, J. S.; Mayor, U.; Freund, S. M. V; Fersht, A. R. Solution Structure of a Protein Denatured State and Folding Intermediate. *Nature* **2005**, *437* (7061), 1053−1056.
(291) 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.
(292) Hamada, D.; Segawa, S.; Goto, Y. Non-Native α-Helical Intermediate in the Refolding of β-Lactoglobulin, a Predominantly β-Sheet Protein. *Nat. Struct. Mol. Biol.* **1996**, *3* (10), 868−873.
(293) Rothwarf, D. M.; Scheraga, H. A. Role of Non-Native Aromatic and Hydrophobic Interactions in the Folding of Hen Egg White Lysozyme. *Biochemistry* **1996**, *35* (43), 13797−13807.
(294) Konuma, T.; Sakurai, K.; Yagi, M.; Goto, Y.; Fujisawa, T.; Takahashi, S. Highly Collapsed Conformation of the Initial Folding Intermediates of β -Lactoglobulin with Non-Native α-Helix. *J. Mol. Biol.* **2015**, *427*, 3158.
(295) Rosen, L. E.; Connell, K. B.; Marqusee, S. Evidence for Close Side-Chain Packing in an Early Protein Folding Intermediate Previously Assumed to Be a Molten Globule. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (41), 14746−14751.
(296) Rosen, L. E.; Kathuria, S. V.; Matthews, C. R.; Bilsel, O.; Marqusee, S. Non-Native Structure Appears in Microseconds during the Folding of E. Coli RNase H. *J. Mol. Biol.* **2015**, *427* (2), 443−453.
(297) Zaidi, F. N.; Nath, U.; Udgaonkar, J. B. Multiple Intermediates and Transition States during Protein Unfolding. *Nat. Struct. Biol.* **1997**, *4* (12), 1016−1024.
(298) Kauzmann, W. Some Factors in the Interpretation of Protein Denaturation. *Adv. Protein Chem.* **1959**, *14*, 1−63.
(299) Baldwin, R. L. The New View of Hydrophobic Free Energy. *FEBS Lett.* **2013**, *587* (8), 1062−1066.
(300) Frank, H. S.; Evans, M. W. Free Volume and Entropy in Condensed Systems III. Entropy in Binary Liquid Mixtures; Partial Molal Entropy in Dilute Solutions; Structure and Thermodynamics in Aqueous Electrolytes. *J. Chem. Phys.* **1945**, *13* (11), 507−532.
(301) Baldwin, R. L.; Rose, G. D. How the Hydrophobic Factor Drives Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (44), 12462−12466.
(302) Jorgensen, W. L.; Gao, J.; Ravimohan, C. Monte Carlo Simulations of Alkanes in Water: hydration number and the hydrophobic effect. *J. Phys. Chem.* **1985**, *89*, 3470−3473.
(303) Southall, N. T.; Dill, K. A.; Haymet, A. D. J. A View of the Hydrophobic Effect. *J. Phys. Chem. B* **2002**, *106* (3), 521−533.
(304) Lee, B. A Procedure for Calculating Thermodynamic Functions of Cavity Formation from the Pure Solvent Simulation Data. *J. Chem. Phys.* **1985**, *83* (5), 2421−2425.

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

(305) Lee, B. Solvent Reorganization Contribution to the Transfer Thermodynamics of Small Nonpolar Molecules. *Biopolymers* **1991**, *31* (8), 993−1008.
(306) Merlino, A.; Pontillo, N.; Graziano, G. A Driving Force for Polypeptide and Protein Collapse. *Phys. Chem. Chem. Phys.* **2017**, *19* (1), 751−756.
(307) Baldwin, R. L. The Nature of Protein Folding Pathways: The Classical versus the New View. *J. Biomol. NMR* **1995**, *5* (2), 103−109.
(308) Engen, J. R.; Smithgall, T. E.; Gmeiner, W. H.; Smith, D. L. Identification and Localization of Slow, Natural, Cooperative Unfolding in the Hematopoietic Cell Kinase SH3 Domain by Amide Hydrogen Exchange and Mass Spectrometry. *Biochemistry* **1997**, *36* (47), 14384−14391.
(309) Arrington, C. B.; Robertson, A. D. Kinetics and Thermodynamics of Conformational Equilibria in Native Proteins by Hydrogen Exchange. *Methods Enzymol.* **2000**, *323*, 104−124.
(310) Wani, A. H.; Udgaonkar, J. B. HX-ESI-MS and Optical Studies of the Unfolding of Thioredoxin Indicate Stabilization of a Partially Unfolded, Aggregation-Competent Intermediate at Low pH. *Biochemistry* **2006**, *45* (37), 11226−11238.
(311) Fersht, A. R. Characterizing Transition States in Protein Folding: An Essential Step in the Puzzle. *Curr. Opin. Struct. Biol.* **1995**, *5* (1), 79−84.
(312) Udgaonkar, J. B.; Baldwin, R. L. Early Folding Intermediate of Ribonuclease A. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87* (21), 8197−8201.
(313) Hu, W.; Walters, B. T.; Kan, Z. Y.; Mayne, L.; Rosen, L. E.; Marqusee, S.; Englander, S. W. Stepwise Protein Folding at near Amino Acid Resolution by Hydrogen Exchange and Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (19), 7684−7689.
(314) Aghera, N.; Udgaonkar, J. B. Stepwise Assembly of β-Sheet Structure during the Folding of an SH3 Domain Revealed by a Pulsed Hydrogen Exchange Mass Spectrometry Study. *Biochemistry* **2017**, *56* (29), 3754−3769.
(315) Aksel, T.; Barrick, D. Direct Observation of Parallel Folding Pathways Revealed Using a Symmetric Repeat Protein System. *Biophys. J.* **2014**, *107* (1), 220−232.
(316) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis. *Proteins Struct. Funct. Bioinforma.* **1995**, *21* (3), 167−195.
(317) Wolynes, P.; Onuchic, J.; Thirumalai, D. Navigating the Folding Routes. *Science.* **1995**, *267* (5204), 1619−1620.
(318) Dill, K. A.; Chan, H. S. From Levinthal to Pathways to Funnels. *Nat. Struct. Mol. Biol.* **1997**, *4* (1), 10−19.
(319) Zwanzig, R. Two-State Models of Protein Folding Kinetics. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (1), 148−150.
(320) Ozkan, S. B.; Dill, K. A.; Bahar, I. Fast-Folding Protein Kinetics, Hidden Intermediates, and the Sequential Stabilization Model. *Protein Sci.* **2002**, *11* (8), 1958−1970.
(321) Ozkan, S. B.; Wu, G. A.; Chodera, J. D.; Dill, K. A. Protein Folding by Zipping and Assembly. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (29), 11987−11992.
(322) Harrison, S. C.; Durbin, R. Is There a Single Pathway for the Folding of a Polypeptide Chain? *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82* (12), 4028−4030.
(323) Lindberg, M.; Tångrot, J.; Oliveberg, M. Complete Change of the Protein Folding Transition State upon Circular Permutation. *Nat. Struct. Biol.* **2002**, *9* (11), 818−822.
(324) Krantz, B. A.; Dothager, R. S.; Sosnick, T. R. Discerning the Structure and Energy of Multiple Transition States in Protein Folding Using ψ-Analysis. *J. Mol. Biol.* **2004**, *337* (2), 463−475.
(325) Jha, S. K.; Dasgupta, A.; Malhotra, P.; Udgaonkar, J. B. Identification of Multiple Folding Pathways of Monellin Using Pulsed Thiol Labeling and Mass Spectrometry. *Biochemistry* **2011**, *50* (15), 3062−3074.
(326) Wirth, A. J.; Liu, Y.; Prigozhin, M. B.; Schulten, K.; Gruebele, M. Comparing Fast Pressure Jump and Temperature Jump Protein Folding Experiments and Simulations. *J. Am. Chem. Soc.* **2015**, *137* (22), 7152−7159.
(327) Zhuravlev, P. I.; Hinczewski, M.; Chakrabarti, S.; Marqusee, S.; Thirumalai, D. Force-Dependent Switch in Protein Unfolding Pathways and Transition-State Movements. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (6), E715−E724.
(328) Sen Mojumdar, S.; Scholl, Z. N.; Dee, D. R.; Rouleau, L.; Anand, U.; Garen, C.; Woodside, M. T. Partially Native Intermediates Mediate Misfolding of SOD1 in Single-Molecule Folding Trajectories. *Nat. Commun.* **2017**, *8* (1), 1881.
(329) Charlier, C.; Reid Alderson, T.; Courtney, J. M.; Ying, J.; Anfinrud, P.; Bax, A. Study of Protein Folding under Native Conditions by Rapidly Switching the Hydrostatic Pressure inside an NMR Sample Cell. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (18), E4169−E4178.
(330) Radford, S. E.; Dobson, C. M.; Evans, P. A. The Folding of Hen Lysozyme Involves Partially Structured Intermediates and Multiple Pathways. *Nature* **1992**, *358* (6384), 302−307.
(331) Shastry, M. C. R.; Udgaonkar, J. B. The Folding Mechanism of Barstar: Evidence for Multiple Pathways and Multiple Intermediates. *J. Mol. Biol.* **1995**, *247* (5), 1013−1027.
(332) Goldbeck, R. A.; Thomas, Y. G.; Chen, E.; Esquerra, R. M.; Kliger, D. S. Multiple Pathways on a Protein-Folding Energy Landscape: Kinetic Evidence. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (6), 2782−2787.
(333) Bieri, O.; Wildegger, G.; Bachmann, A.; Wagner, C.; Kiefhaber, T. A Salt-Induced Kinetic Intermediate Is on a New Parallel Pathway of Lysozyme Folding. *Biochemistry* **1999**, *38* (38), 12460−12470.
(334) Rami, B. R.; Udgaonkar, J. B. PH-Jump-Induced Folding and Unfolding Studies of Barstar: Evidence for Multiple Folding and Unfolding Pathways. *Biochemistry* **2001**, *40* (50), 15267−15279.
(335) Aghera, N.; Udgaonkar, J. B. The Utilization of Competing Unfolding Pathways of Monellin Is Dictated by Enthalpic Barriers. *Biochemistry* **2013**, *52* (34), 5770−5779.
(336) Aggarwal, V.; Kulothungan, S. R.; Balamurali, M. M.; Saranya, S. R.; Varadarajan, R.; Ainavarapu, S. R. K. Ligand-Modulated Parallel Mechanical Unfolding Pathways of Maltose-Binding Proteins. *J. Biol. Chem.* **2011**, *286* (32), 28056−28065.
(337) Wright, C. F.; Lindorff-Larsen, K.; Randles, L. G.; Clarke, J. Parallel Protein-Unfolding Pathways Revealed and Mapped. *Nat. Struct. Biol.* **2003**, *10* (8), 658−662.
(338) Moulick, R.; Goluguri, R. R.; Udgaonkar, J. B. Ruggedness in the Free Energy Landscape Dictates Misfolding of the Prion Protein. *J. Mol. Biol.* **2019**, *431* (4), 807−824.
(339) Yoo, J.; Louis, J. M.; Chung, H. S. Diverse Folding Pathways of HIV-1 Protease Monomer on a Rugged Energy Landscape. *Biophys. J.* **2019**, *117* (8), 1456−1466.
(340) Satija, R.; Berezhkovskii, A. M.; Makarov, D. E. Broad Distributions of Transition-Path Times Are Fingerprints of Multidimensionality of the Underlying Free Energy Landscapes. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (44), 27116−27123.
(341) Viguera, A. R.; Blanco, F. J.; Serrano, L. The Order of Secondary Structure Elements Does Not Determine the Structure of a Protein but Does Affect Its Folding Kinetics. *J. Mol. Biol.* **1995**, *247* (4), 670−681.
(342) Leeson, D. T.; Gai, F.; Rodriguez, H. M.; Gregoret, L. M.; Dyer, R. B. Protein Folding and Unfolding on a Complex Energy Landscape. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (6), 2527−2532.
(343) Hutton, R. D.; Wilkinson, J.; Faccin, M.; Sivertsson, E. M.; Pelizzola, A.; Lowe, A. R.; Bruscolini, P.; Itzhaki, L. S. Mapping the Topography of a Protein Energy Landscape. *J. Am. Chem. Soc.* **2015**, *137* (46), 14610−14625.
(344) Narayan, A.; Naganathan, A. N. Evidence for the Sequential Folding Mechanism in RNase H from an Ensemble-Based Model. *J. Phys. Chem. B* **2014**, *118* (19), 5050−5058.
(345) Gopi, S.; Singh, A.; Suresh, S.; Paul, S.; Ranu, S.; Naganathan, A. N. Toward a Quantitative Description of Microscopic Pathway Heterogeneity in Protein Folding. *Phys. Chem. Chem. Phys.* **2017**, *19* (31), 20891−20903.
(346) Gershenson, A.; Gosavi, S.; Faccioli, P.; Wintrode, P. L. Successes and Challenges in Simulating the Folding of Large Proteins. *J. Biol. Chem.* **2020**, *295* (1), 15−33.

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

(347) Neelamraju, S.; Wales, D. J.; Gosavi, S. Protein Energy Landscape Exploration with Structure-Based Models. *Curr. Opin. Struct. Biol.* **2020**, *64*, 145−151.
(348) Borreguero, J. M.; Ding, F.; Buldyrev, S. V.; Stanley, H. E.; Dokholyan, N. V. Multiple Folding Pathways of the SH3 Domain. *Biophys. J.* **2004**, *87* (1), 521−533.
(349) Andrec, M.; Felts, A. K.; Gallicchio, E.; Levy, R. M. Protein Folding Pathways from Replica Exchange Simulations and a Kinetic Network Model. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6801−6806.
(350) Lam, A. R.; Borreguero, J. M.; Ding, F.; Dokholyan, N. V.; Buldyrev, S. V.; Stanley, H. E.; Shakhnovich, E. Parallel Folding Pathways in the SH3 Domain Protein. *J. Mol. Biol.* **2007**, *373* (5), 1348−1360.
(351) Noé, F.; Schütte, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. Constructing the Equilibrium Ensemble of Folding Pathways from Short Off-Equilibrium Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (45), 19011−19016.
(352) Beccara, S. a.; Skrbic, T.; Covino, R.; Faccioli, P. Dominant Folding Pathways of a WW Domain. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (7), 2330−2335.
(353) Prigozhin, M. B.; Gruebele, M. Microsecond Folding Experiments and Simulations: A Match Is Made. *Phys. Chem. Chem. Phys.* **2013**, *15* (10), 3372.
(354) Rizzuti, B.; Daggett, V. Using Simulations to Provide the Framework for Experimental Protein Folding Studies. *Arch. Biochem. Biophys.* **2013**, *531* (1−2), 128−135.
(355) Lapidus, L. J.; Acharya, S.; Schwantes, C. R.; Wu, L.; Shukla, D.; King, M.; DeCamp, S. J.; Pande, V. S. Complex Pathways in Folding of Protein G Explored by Simulation and Experiment. *Biophys. J.* **2014**, *107* (4), 947−955.
(356) Zanetti-Polzi, L.; Davis, C. M.; Gruebele, M.; Dyer, R. B.; Amadei, A.; Daidone, I. Parallel Folding Pathways of Fip35 WW Domain Explained by Infrared Spectra and Their Computer Simulation. *FEBS Lett.* **2017**, *591* (20), 3265−3275.
(357) Davis, C. M.; Zanetti-Polzi, L.; Gruebele, M.; Amadei, A.; Dyer, R. B.; Daidone, I. A Quantitative Connection of Experimental and Simulated Folding Landscapes by Vibrational Spectroscopy. *Chem. Sci.* **2018**, *9* (48), 9002−9011.
(358) Kantaev, R.; Riven, I.; Goldenzweig, A.; Barak, Y.; Dym, O.; Peleg, Y.; Albeck, S.; Fleishman, S. J.; Haran, G. Manipulating the Folding Landscape of a Multidomain Protein. *J. Phys. Chem. B* **2018**, *122* (49), 11030−11038.
(359) Best, R. B.; Hummer, G. Reaction Coordinates and Rates from Transition Paths. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6732−6737.
(360) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; et al. Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science (80-.).* **2010**, *330* (6002), 341−346.
(361) Beauchamp, K. A.; McGibbon, R.; Lin, Y.-S.; Pande, V. S. Simple Few-State Models Reveal Hidden Complexity in Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (44), 17807−17813.
(362) Olofsson, M.; Hansson, S.; Hedberg, L.; Logan, D. T.; Oliveberg, M. Folding of S6 Structures with Divergent Amino Acid Composition: Pathway Flexibility Within Partly Overlapping Foldons. *J. Mol. Biol.* **2007**, *365* (1), 237−248.
(363) Connell, K. B.; Miller, E. J.; Marqusee, S. The Folding Trajectory of RNase H Is Dominated by Its Topology and Not Local Stability: A Protein Engineering Study of Variants That Fold via Two-State and Three-State Mechanisms. *J. Mol. Biol.* **2009**, *391* (2), 450−460.
(364) Pradeep, L.; Udgaonkar, J. B. Differential Salt-Induced Stabilization of Structure in the Initial Folding Intermediate Ensemble of Barstar. *J. Mol. Biol.* **2002**, *324* (2), 331−347.
(365) Hu, W.; Kan, Z. Y.; Mayne, L.; Englander, S. W. Cytochrome c Folds through Foldon-Dependent Native-like Intermediates in an Ordered Pathway. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (14), 3809−3814.
(366) Baldwin, R. L. Clash between Energy Landscape Theory and Foldon-Dependent Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (32), 8442−8443.
(367) Bédard, S.; Krishna, M. M. G.; Mayne, L.; Englander, S. W. Protein Folding: Independent Unrelated Pathways or Predetermined Pathway with Optional Errors. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (20), 7182−7187.
(368) Arrington, C. B.; Teesch, L. M.; Robertson, A. D. Defining Protein Ensembles with Native-State NH Exchange: Kinetics of Interconversion and Cooperative Units from Combined NMR and MS Analysis. *J. Mol. Biol.* **1999**, *285* (3), 1265−1275.
(369) Malhotra, P.; Udgaonkar, J. B. Tuning Cooperativity on the Free Energy Landscape of Protein Folding. *Biochemistry* **2015**, *54* (22), 3431−3441.
(370) Jethva, P. N.; Udgaonkar, J. B. Modulation of the Extent of Cooperative Structural Change during Protein Folding by Chemical Denaturant. *J. Phys. Chem. B* **2017**, *121* (35), 8263−8275.
(371) Chung, H. S.; McHale, K.; Louis, J. M.; Eaton, W. A. Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times. *Science.* **2012**, *335* (6071), 981−984.
(372) Chung, H. S.; Eaton, W. A. Protein Folding Transition Path Times from Single Molecule FRET. *Curr. Opin. Struct. Biol.* **2018**, *48*, 30−39.
(373) Eaton, W. A. Modern Kinetics and Mechanism of Protein Folding: A Retrospective. *J. Phys. Chem. B* **2021**, *125* (14), 3452−3467.
(374) Yu, H.; Liu, X.; Neupane, K.; Gupta, A. N.; Brigley, A. M.; Solanki, A.; Sosova, I.; Woodside, M. T. Direct Observation of Multiple Misfolding Pathways in a Single Prion Protein Molecule. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (14), 5283−5288.
(375) Neupane, K.; Manuel, A. P.; Woodside, M. T. Protein Folding Trajectories Can Be Described Quantitatively by One-Dimensional Diffusion over Measured Energy Landscapes. *Nat. Phys.* **2016**, *12* (7), 700−703.
(376) Petrosyan, R.; Patra, S.; Rezajooei, N.; Garen, C. R.; Woodside, M. T. Unfolded and Intermediate States of PrP Play a Key Role in the Mechanism of Action of an Antiprion Chaperone. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118* (9), e2010213118.
(377) Hoffer, N. Q.; Woodside, M. T. Probing Microscopic Conformational Dynamics in Folding Reactions by Measuring Transition Paths. *Curr. Opin. Chem. Biol.* **2019**, *53*, 68−74.
(378) Pirchi, M.; Ziv, G.; Riven, I.; Cohen, S. S.; Zohar, N.; Barak, Y.; Haran, G. Single-Molecule Fluorescence Spectroscopy Maps the Folding Landscape of a Large Protein. *Nat. Commun.* **2011**, *2* (1), 493−497.
(379) Stigler, J.; Ziegler, F.; Gieseke, A.; Gebhardt, J. C. M.; Rief, M. The Complex Folding Network of Single Calmodulin Molecules. *Science.* **2011**, *334* (6055), 512−516.
(380) Jagannathan, B.; Elms, P. J.; Bustamante, C.; Marqusee, S. Direct Observation of a Force-Induced Switch in the Anisotropic Mechanical Unfolding Pathway of a Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (44), 17820−17825.
(381) Kotamarthi, H. C.; Sharma, R.; Narayan, S.; Ray, S.; Ainavarapu, S. R. K. Multiple Unfolding Pathways of Leucine Binding Protein (LBP) Probed by Single-Molecule Force Spectroscopy (SMFS). *J. Am. Chem. Soc.* **2013**, *135* (39), 14768−14774.
(382) Neupane, K.; Solanki, A.; Sosova, I.; Belov, M.; Woodside, M. T. Diverse Metastable Structures Formed by Small Oligomers of α-Synuclein Probed by Force Spectroscopy. *PLoS One* **2014**, *9* (1), e86495.
(383) Solanki, A.; Neupane, K.; Woodside, M. T. Single-Molecule Force Spectroscopy of Rapidly Fluctuating, Marginally Stable Structures in the Intrinsically Disordered Protein α-Synuclein. *Phys. Rev. Lett.* **2014**, *112* (15), 158103.
(384) Yu, H.; Dee, D. R.; Liu, X.; Brigley, A. M.; Sosova, I.; Woodside, M. T. Protein Misfolding Occurs by Slow Diffusion across Multiple Barriers in a Rough Energy Landscape. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (27), 8308−8313.

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX

(385) Yan, Z.; Wang, J. Funneled Energy Landscape Unifies Principles of Protein Binding and Evolution. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (44), 27218−27223.
(386) Llinas, M.; Gillespie, B.; Dahlquist, F. W.; Marqusee, S. The Energetics of T4 Lysozyme Reveal a Hierarchy of Conformations. *Nat. Struct. Biol.* **1999**, *6* (11), 1072−1078.
(387) Jackson, S. E.; Fersht, A. R. Folding of Chymotrypsin Inhibitor 2. 1. Evidence for a Two-State Transition. *Biochemistry* **1991**, *30* (43), 10428−10435.
(388) Garcia-Mira, M. M.; Sadqi, M.; Fischer, N.; Sanchez-Ruiz, J. M.; Munoz, V. Experimental Identification of Downhill Protein Folding. *Science (80-.).* **2002**, *298* (5601), 2191−2195.
(389) Sadqi, M.; Fushman, D.; Muñoz, V. Atom-by-Atom Analysis of Global Downhill Protein Folding. *Nature* **2006**, *442* (7100), 317−321.
(390) Fung, A.; Li, P.; Godoy-Ruiz, R.; Sanchez-Ruiz, J. M.; Muñoz, V. Expanding the Realm of Ultrafast Protein Folding: GpW, a Midsize Natural Single-Domain with A+β Topology That Folds Downhill. *J. Am. Chem. Soc.* **2008**, *130* (23), 7489−7495.
(391) Sborgi, L.; Verma, A.; Piana, S.; Lindorff-Larsen, K.; Cerminara, M.; Santiveri, C. M.; Shaw, D. E.; De Alba, E.; Muñoz, V. Interaction Networks in Protein Folding via Atomic-Resolution Experiments and Long-Time-Scale Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2015**, *137* (20), 6506−6516 jacs.5b02324..
(392) Lindhoud, S.; Pirchi, M.; Westphal, A. H.; Haran, G.; Van Mierlo, C. P. M. Gradual Folding of an Off-Pathway Molten Globule Detected at the Single-Molecule Level. *J. Mol. Biol.* **2015**, *427* (19), 3148−3157.
(393) Jha, S. K.; Udgaonkar, J. B. Free Energy Barriers in Protein Folding and Unfolding Reactions. *Curr. Sci.* **2010**, 457−475.
(394) Oikawa, H.; Suzuki, Y.; Saito, M.; Kamagata, K.; Arai, M.; Takahashi, S. Microsecond Dynamics of an Unfolded Protein by a Line Confocal Tracking of Single Molecule Fluorescence. *Sci. Rep.* **2013**, *3*, 1−6.
(395) Schuler, B.; Hofmann, H. Single-Molecule Spectroscopy of Protein Folding Dynamics — Expanding Scope and Timescales. *Curr. Opin. Struct. Biol.* **2013**, *23* (1), 36−47.
(396) Holtzer, M.E.; Lovett, E.G.; d'Avignon, D.A.; Holtzer, A. Thermal Unfolding in a GCN4-like Leucine Zipper: 13C Alpha NMR Chemical Shifts and Local Unfolding Curves. *Biophys J.* **1997**, *73* (2), 1031−1041.
(397) Mohan, P. M. K.; Chakraborty, S.; Hosur, R. V. NMR Investigations on Residue Level Unfolding Thermodynamics in DLC8 Dimer by Temperature Dependent Native State Hydrogen Exchange. *J. Biomol. NMR* **2009**, *44* (1), 1−11.
(398) Sridevi, K.; Juneja, J.; Bhuyan, A. K.; Krishnamoorthy, G.; Udgaonkar, J. B. The Slow Folding Reaction of Barstar: The Core Tryptophan Region Attains Tight Packing before Substantial Secondary and Tertiary Structure Formation and Final Compaction of the Polypeptide Chain. *J. Mol. Biol.* **2000**, *302* (2), 479−495.
(399) Rami, B. R.; Udgaonkar, J. B. Mechanism of Formation of a Productive Molten Globule Form of Barstar. *Biochemistry* **2002**, *41* (6), 1710−1716.
(400) Roder, H.; Elöve, G. A.; Englander, S. W. Structural Characterization of Folding Intermediates in Cytochrome c by H-Exchange Labelling and Proton NMR. *Nature* **1988**, *335* (6192), 700−704.
(401) Jha, S. K.; Udgaonkar, J. B. Direct Evidence for a Dry Molten Globule Intermediate during the Unfolding of a Small Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (30), 12289−12294.
(402) Sarkar, S. S.; Udgaonkar, J. B.; Krishnamoorthy, G. Unfolding of a Small Protein Proceeds via Dry and Wet Globules and a Solvated Transition State. *Biophys. J.* **2013**, *105* (10), 2392−2402.
(403) Jha, S. K.; Marqusee, S. Kinetic Evidence for a Two-Stage Mechanism of Protein Denaturation by Guanidinium Chloride. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (13), 4856−4861.
(404) Acharya, N.; Mishra, P.; Jha, S. K. Evidence for Dry Molten Globule-Like Domains in the PH-Induced Equilibrium Folding Intermediate of a Multidomain Protein. *J. Phys. Chem. Lett.* **2016**, *7* (1), 173−179.
(405) Kuwajima, K. The Molten Globule State as a Clue for Understanding the Folding and Cooperativity of Globular-Protein Structure. *Proteins Struct. Funct. Genet.* **1989**, *6* (2), 87−103.
(406) Georgescu, R. E.; Li, J.-H.; Goldberg, M. E.; Tasayco, M. L.; Chaffotte, A. F. Proline Isomerization-Independent Accumulation of an Early Intermediate and Heterogeneity of the Folding Pathways of a Mixed α/β Protein, Escherichia Coli Thioredoxin †. *Biochemistry* **1998**, *37* (28), 10286−10297.
(407) Sosnick, T. R.; Shtilerman, M. D.; Mayne, L.; Englander, S. W. Ultrafast Signals in Protein Folding and the Polypeptide Contracted State. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (16), 8545−8550.
(408) Parker, M. J.; Marqusee, S. The Cooperativity of Burst Phase Reactions Explored. *J. Mol. Biol.* **1999**, *293* (5), 1195−1210.
(409) Hagen, S. J. Probe-Dependent and Nonexponential Relaxation Kinetics: Unreliable Signatures of Downhill Protein Folding. *Proteins Struct. Funct. Bioinforma.* **2007**, *68* (1), 205−217.
(410) Stigter, D.; Alonso, D. O.; Dill, K. A. Protein Stability: Electrostatics and Compact Denatured States. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88* (10), 4176−4180.
(411) Sabelko, J.; Ervin, J.; Gruebele, M. Observation of Strange Kinetics in Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (11), 6031−6036.
(412) Liu, F.; Gruebele, M. Tuning λ 6−85 Towards Downhill Folding at Its Melting Temperature. *J. Mol. Biol.* **2007**, *370* (3), 574−584.
(413) Gruebele, M. Downhill Protein Folding: Evolution Meets Physics. *Comptes Rendus - Biol.* **2005**, *328* (8), 701−712.
(414) Malhotra, P.; Jethva, P. N.; Udgaonkar, J. B. Chemical Denaturants Smoothen Ruggedness on the Free Energy Landscape of Protein Folding. *Biochemistry* **2017**, *56* (31), 4053−4063.
(415) Spudich, G. M.; Miller, E. J.; Marqusee, S. Destabilization of the Escherichia Coli RNase H Kinetic Intermediate: Switching between a Two-State and Three-State Folding Mechanism. *J. Mol. Biol.* **2004**, *335* (2), 609−618.
(416) Kiefhaber, T.; Bachmann, A.; Wildegger, G.; Wagner, C. Direct Measurement of Nucleation and Growth Rates in Lysozyme Folding †. *Biochemistry* **1997**, *36* (17), 5108−5112.
(417) Wolynes, P. G.; Eaton, W. A.; Fersht, A. R. Chemical Physics of Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (44), 17770−17771.
(418) Onuchic, J. N.; Wolynes, P. G. Theory of Protein Folding. *Curr. Opin. Struct. Biol.* **2004**, *14* (1), 70−75.
(419) Go, N. Theoretical Studies of Protein Folding. *Annu. Rev. Biophys. Bioeng.* **1983**, *12* (1), 183−210.
(420) Chen, T.; Chan, H. S. Effects of Desolvation Barriers and Sidechains on Local-Nonlocal Coupling and Chevron Behaviors in Coarse-Grained Models of Protein Folding. *Phys. Chem. Chem. Phys.* **2014**, *16* (14), 6460−6479.
(421) Poland, D.; Scheraga, H. A. Phase Transitions in One Dimension and the Helix—Coil Transition in Polyamino Acids. *J. Chem. Phys.* **1966**, *45* (5), 1456−1463.
(422) Thompson, P. A.; Eaton, W. A.; Hofrichter, J. Laser Temperature Jump Study of the Helix⇌Coil Kinetics of an Alanine Peptide Interpreted with a 'Kinetic Zipper" Model.'. *Biochemistry* **1997**, *36* (30), 9200−9210.
(423) Gelman, H.; Gruebele, M. Fast Protein Folding Kinetics. *Q. Rev. Biophys.* **2014**, *47* (2), 95−142.

https://doi.org/10.1021/acs.chemrev.1c00704
Chem. Rev. XXXX, XXX, XXX−XXX