doi:10.1016/j.jmb.2012.01.016 J. Mol. Biol. (2012) 418, 226–236

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Small-Angle X-ray Scattering and Single-Molecule FRET Spectroscopy Produce Highly Divergent Views of the Low-Denaturant Unfolded State

Tae Yeon Yoo [1]†, Steve P. Meisburger [2]†, James Hinshaw [3]†, Lois Pollack [2]*, Gilad Haran [4]*, Tobin R. Sosnick [1,5,6]* and Kevin Plaxco [7,8]*

[1]*Department of Physics, University of Chicago, Chicago, IL 60637, USA*
[2]*School of Applied and Engineering Physics, Cornell University, Ithaca, NY 14853, USA*
[3]*Department of Chemistry, University of Chicago, Chicago, IL 60637, USA*
[4]*Chemical Physics Department, Weizmann Institute of Science, Rehovot 76100, Israel*
[5]*Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637, USA*
[6]*Institute for Biophysical Dynamics, University of Chicago, Chicago, IL 60637, USA*
[7]*Department of Chemistry and Biochemistry, University of California, Santa Barbara, Santa Barbara, CA 93106, USA*
[8]*Interdepartmental Program in Biomolecular Science and Engineering, University of California, Santa Barbara, Santa Barbara, CA 93106, USA*

Received 3 October 2011;
received in revised form
22 December 2011;
accepted 18 January 2012
Available online
27 January 2012

Edited by M. Yaniv

Keywords:
physiological unfolded state;
hydrophobic collapse;
nonspecific;
coil-to-globule

The results of more than a dozen single-molecule Förster resonance energy transfer (smFRET) experiments suggest that chemically unfolded polypeptides invariably collapse from an expanded random coil to more compact dimensions as the denaturant concentration is reduced. In sharp contrast, small-angle X-ray scattering (SAXS) studies suggest that, at least for single-domain proteins at non-zero denaturant concentrations, such compaction may be rare. Here, we explore this discrepancy by studying protein L, a protein previously studied by SAXS (at 5 °C), which suggested fixed unfolded-state dimensions from 1.4 to 5 M guanidine hydrochloride (GuHCl), and by smFRET (at 25 °C), which suggested that, in contrast, the chain contracts by 15–30% over this same denaturant range. Repeating the earlier SAXS study under the same conditions employed in the smFRET studies, we observe little, if any, evidence that the unfolded state of protein L contracts as the concentration of GuHCl is reduced. For example, scattering profiles (and thus the shape and dimensions) collected within ~4 ms after dilution to as low as 0.67 M GuHCl are effectively indistinguishable from those observed at equilibrium at higher denaturant. Our results thus argue that the disagreement between SAXS and smFRET is statistically significant and that the experimental evidence in favor of obligate polypeptide collapse at low denaturant cannot be considered conclusive yet.

**Corresponding authors.* E-mail address: kwp@chem.ucsb.edu.

† T.Y.Y., S.P.M., and J.H. contributed equally to this work.

Abbreviations used: smFRET, single-molecule Förster resonance energy transfer; SAXS, small-angle X-ray scattering; GuHCl, guanidine hydrochloride; APS, Advanced Photon Source; CHESS, Cornell High Energy Synchrotron Source.

Introduction

Imagine you have in your test tube a simple, single-domain protein lacking disulfide bonds or other cross-links. If you were to unfold this protein at high levels of chemical denaturant (e.g., above 6 M guanidine hydrochloride, GuHCl), it would likely expand to the dimensions expected for an unstructured, random-coil homopolymer (e.g., Refs. 1 and 2). If you were to then gradually reduce the concentration of denaturant, more and more of the molecules in your test tube would fold. But what happens to the (increasingly poorly populated) set of molecules that remain unfolded? Do they also contract as the denaturant concentration is reduced, or do they remain expanded across a broad range of denaturant concentrations? Resolution of this simple—if fundamental—question in protein physics, a question that speaks directly to both the thermodynamics and kinetics of protein folding, remains mired in controversy. Specifically, the two experimental approaches that perhaps most directly and quantitatively speak to this question, single-molecule Förster resonance energy transfer (smFRET) and small-angle X-ray scattering (SAXS), produce highly discordant answers to this seemingly simple question.

The controversy we are addressing revolves around the behavior of the unfolded states of single-domain proteins in response to changing levels of chemical denaturant.[3] To date, at least a dozen reported studies have employed smFRET to monitor such unfolded states, the results of which are universally interpreted in terms of a significant, monotonic contraction of the unfolded state as the concentration of chemical denaturant is reduced (e.g., Refs. 4–15). Corroboration of these studies is provided by ensemble FRET studies, including, for example, equilibrium FRET studies of simple $(Gly–Ser)_N$ polymers[16] and time-resolved FRET studies of unfolded protein L prior to refolding,[17] both of which also suggest that their respective unfolded states contract significantly as the denaturant concentration is reduced. In clear contradiction to this picture, however, equilibrium SAXS experiments find that the dimensions of the unfolded baseline observed in equilibrium chemical melts are almost always independent of denaturant concentration (e.g., Refs. 18–27; for a rare potential counterexample, see Ref. 18). A series of time-resolved stopped-flow SAXS experiments likewise suggest that, for many single-domain proteins, the dimensions of the transient state formed upon rapid dilution to low denaturant are also indistinguishable from those seen at higher denaturant.[22,24,27] A significant disagreement thus exists between SAXS, which, for many proteins, fails to "see" significant contraction of the unfolded state as the level of denaturant is reduced, and smFRET studies, which universally suggest a significant, steady contraction with decreasing denaturant.

The difference in the unfolded state behaviors observed by smFRET and by SAXS is both real and scientifically significant. That is, although much of the discrepancy could simply reflect variations in the behavior of different proteins, protein L, the single protein that has been investigated by both SAXS and smFRET, produces highly discordant results via the two approaches (Fig. 1). In the first study of this protein, Baker *et al.* employed SAXS to argue that the unfolded-state dimensions of protein L are invariant over a wide range of GuHCl concentrations.[22] Specifically, they reported that, at 27.1 ± 1.6 Å, the radius of gyration (R_g) of the unfolded ensemble of protein L transiently populated prior to refolding in 1.4 M GuHCl is experimentally indistinguishable from the 26.0 ± 0.3 Å R_g observed at equilibrium in 5 M GuHCl. More recently, however, Sherman and Haran used smFRET to argue that the R_g of unfolded protein L contracts from ~24.5 Å to 18 Å as the denaturant concentration is reduced over this same

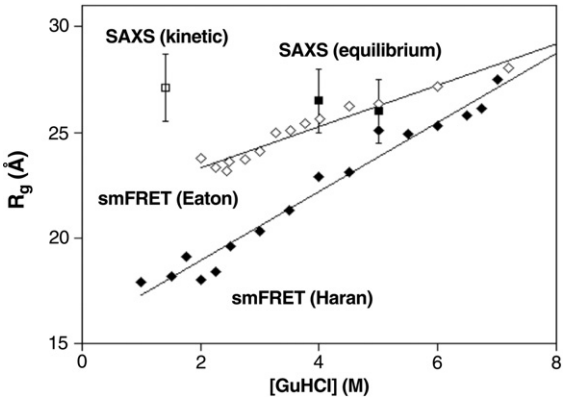

Fig. 1. The dimensions of unfolded protein L have been studied using both smFRET and SAXS. The results of these studies, however, are highly discordant. Specifically, time-resolved (open squares) and equilibrium (filled squares) SAXS studies by Baker *et al.*, conducted at 2.5 °C and 5 °C, respectively, suggest that the radius of gyration of the unfolded ensemble transiently populated prior to refolding in 1.4 M GuHCl is experimentally indistinguishable from the dimensions observed at equilibrium at higher denaturant.[22] In contrast, smFRET studies by Sherman and Haran[8] and Eaton *et al.*[9] suggest that, at least at 20 °C, unfolded protein L contracts significantly upon being transferred from high denaturant to lower denaturant. Here, we explore this discrepancy in greater depth by performing more detailed SAXS studies under solution conditions mimicking those employed in the smFRET studies at both 5 °C and 20 °C. Of note, while the two smFRET data sets differ quantitatively, perhaps due to differences in the parameters used in the data analysis or due to a one-residue difference between the two constructs employed (see analysis in Ref. 28), both argue in favor of significant collapse at low denaturant concentrations.

range.[8] Following this, a similar article by Eaton *et al.* reported contraction from 27 Å at 5 M GuHCl to 24 Å at 2 M.[9] Finally, using a time-resolved ensemble FRET approach, Lapidus *et al.* have observed a large increase in FRET transfer efficiency across unfolded protein L within microseconds of dilution from high denaturant,[17] further suggesting that the unfolded-state ensemble of this protein collapses at low denaturant.

We are thus faced with a significant disagreement. Whereas SAXS suggests that the R_g of unfolded protein L does not change measurably between high and low denaturant, FRET suggests that it undergoes a continuous, readily measurable 15–30% contraction over this span of GuHCl concentrations. This said, even this level of divergence could be due to trivial experimental issues. For example, in order to slow the folding rate of protein L (providing more time for integration in the kinetic experiment), the prior SAXS studies were performed at 2–5 °C, which is well below temperatures at which the equivalent smFRET experiments were conducted. In response, we report here the results of expanded and more detailed SAXS studies of the dimensions of unfolded protein L at 22 °C, allowing us to more firmly establish and better constrain this scientifically important discrepancy.

Results

Two independently measured equilibrium SAXS data sets and one time-resolved kinetic data set were acquired to determine the apparent dimensions of the protein L ensemble as a function of denaturant concentration at pH 7.5, 22 °C. The kinetic data set and one of the two equilibrium data sets were collected at the Advanced Photon Source (APS) and used a minimal, 64-residue construct with a destabilizing N9H substitution chosen in part to facilitate access to lower denaturant conditions (e.g., so that the unfolded state baseline extends to lower GuHCl concentrations). The second equilibrium data set, collected at the Cornell High Energy Synchrotron Source (CHESS), employed a 72-residue construct (retaining a six-histidine affinity tag). The latter was nearly identical with the sequences employed in the smFRET studies of Sherman and Haran[8] and Eaton *et al.*,[9] which differed only in that the latter each contained two cysteine residues (for labeling). The sequences of the two constructs employed here and the three constructs employed previously are presented in Table S1.

Perhaps the most straightforward, assumption-free method of monitoring a change in the shape or dimensions of an unfolded polypeptide is to directly compare X-ray scattering profiles over a range of denaturant concentrations. In doing so, we find little evidence in the equilibrium data sets at denaturant levels well above the protein's folding transition point that the mean dimensions or shape of unfolded protein L change as a function of GuHCl concentration. Specifically, when normalized for differences in overall scattering intensity (due, for example, to changes in beam intensity, sample contrast, denaturant absorbance, or protein concentration), we find that scattering profiles collected between 3.5 and 7.4 M GuHCl are effectively indistinguishable from one another (Fig. 2, top, and Fig. S1). Only at much lower denaturant concentrations—when a significant population of the folded protein is present[31]—do the scattering profiles diverge. Raw scattering profiles thus strongly suggest that, at least at denaturant concentrations well above the unfolding transition, the dimensions of chemically denatured protein L are insensitive to changing denaturant levels.

The above analysis provides a relatively assumption-free method of monitoring changes in molecular dimensions, but does not provide much in the way of quantitative insights with regard to how much contraction might be "hidden" in our data sets. The Guinier approximation (Fig. 2, middle), in contrast, provides a means of determining R_g from scattering profiles and thus allows us to probe this question in more quantitative detail (with the caveat that artifacts at low scattering angles, such as interparticle interference, aggregation, or poor background subtraction, can lead to systematic errors in the estimated radius of gyration). To do so, we first investigate the more comprehensive of our two equilibrium data sets, which was taken at the APS (the CHESS data set was too sparse to fit). This data set covers denaturant concentrations ranging from 1 to 7.4 M GuHCl using a 24-mg/ml, 7.4-M GuHCl stock solution serially diluted with protein- and denaturant-free buffer to final denaturant and protein concentrations of 1 M GuHCl and 3 mg/ml, respectively. Over this range of conditions, we find that the ensemble R_g increases monotonically from 14 to 26 Å (as determined via Guinier analysis of the width of the Gaussian scattering profile at lower scattering angles; e.g., $R_gQ<1.1$–1.3—see Fig. 2, middle) as the population of unfolded protein in the ensemble increases with increasing denaturant concentrations (Fig. 3). The native and unfolded baselines of this equilibrium melt exhibit minimal, if any, dependence on denaturant concentration. Indeed, fits to the standard two-state model,[23,32] which assumes that the slopes of the native and unfolded baselines are zero, are excellent ($R^2=0.972$). Relaxation of this assumption for the unfolded baseline yields a best-fit slope, 0.33 ± 0.35 Å M^{-1}, within error of zero. This said, one subset of the APS equilibrium data set exhibits some baseline slope. That is, although the APS data set does not exhibit any evidence of baseline slope when taken in its entirety, the APS equilibrium SAXS data were collected on

two different days using two different batches of, ostensibly, identical protein. Analysis of the unfolded baselines of these subsets of our data shows that one of them, collected in July 2010, exhibits a statistically significant slope (as determined by an ANOVA test), while the other data set does not show a statistically significant slope (Fig. S3). A comparison of the two data sets suggests, however, that the observed slopes are due to systematic experiment-to-experiment variation rather than due to the physics of unfolded protein L (e.g., the best-fit slopes for the two data sets are of opposite sign and are their difference is statistically significant). Consistent with this, the independently collected CHESS equilibrium data set, while too sparse to constrain a full two-state model, nevertheless fails to produce any significant evidence of baseline slope (Fig. S4).

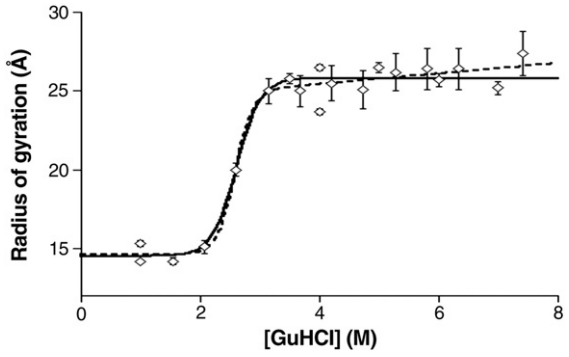

Fig. 3. A fit of the equilibrium SAXS data (APS data at 22 °C; see Fig. S4 for the equivalent CHESS data set) to two-state equilibrium unfolding models does not produce any statistically significant evidence of a sloping unfolded baseline. Specifically, a model lacking baseline slope (continuous line) fits the data well (R^2=0.972), and the addition of a sloped unfolded baseline (broken line) produces only trivial improvement in the residuals (R^2=0.978) and estimates a best-fit slope, 0.33±0.35 Å M^{-1}, within error of zero. In contrast, the equivalent slopes derived using smFRET (Fig. 1) are, at 0.98±0.14 and 1.64±016 Å M^{-1} for the data sets of Eaton and Haran, respectively, highly statistically significant. The confidence intervals on these slopes reflect 95% confidence ranges; the error bars on the data represent standard errors estimated from the fits to Guinier plots.

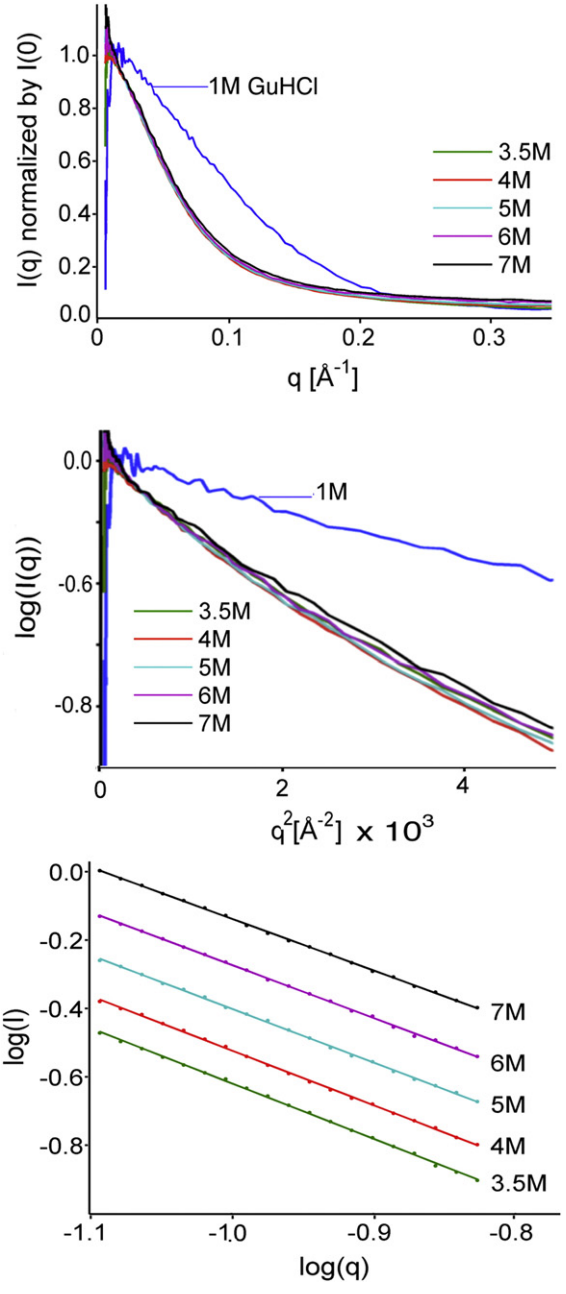

Fig. 2. (Top) The raw scattering profiles [$I(0)$—normalized scattering intensity *versus* q] of protein L collected at equilibrium at high denaturant are effectively indistinguishable over the range 3.5 to 7 M GuHCl, suggesting that the dimensions and shape of the unfolded protein do not change significantly over this range. At 1 M GuHCl, in contrast, the protein folds into its compact native state, dramatically altering its scattering. The profiles presented here are from our APS data set (see Fig. S1 for the equivalent, independently collected CHESS results). CHESS data acquired at both 5 °C and 20 °C are also indistinguishable from one another (see Fig. S2), ruling out temperature-dependent effects. (Middle) Guinier representations of the scattering data likewise suggest that equilibrium unfolded protein L undergoes little, if any, change in dimensions down to at least 3.5 M GuHCl. (Bottom) Finally, analysis of the fractal dimensions[29,30] of unfolded protein L likewise suggests that its unfolded state remains expanded from 7 to 3.5 M GuHCl, as the slopes of log(I) *versus* log (q) plots (corresponding to the fractal dimension, D_m) collected under these conditions all fall between 1.52±0.01 and 1.62±0.01. The observed slopes approximate the value of 1.7 expected for an excluded volume random coil and fall far below the value of 3 expected for a compact globule.

Taken together, these data suggest that the dimensions of unfolded protein L remain fixed above the protein's denaturation transition.

The extent to which the scattering from an unfolded polymer reflects an expanded, random-coil state or a more compact conformation can also be ascertained from the relationship between scattering intensity, I, and scattering angle, q, at larger scattering angles ($q > 1/R_g$), where

$$I(q) \propto q^{-D_m} \quad (1)$$

In this relationship, which holds for many particle types, D_m is a constant that reflects the "fractal dimension" of the scattering particle (see Refs. 29 and 30 and references therein). For polymers, D_m ranges from unity for a fully extended, one-dimensional chain to 3 for a compact globule. The D_m of expanded, unfolded states falls between these two extrema, with an idealized random coil producing a D_m of 2 and an expanded, self-avoiding random-walk polymer producing a D_m of 1.7. Analyzing our protein L data using this approach, we find that, between 3.5 and 7 M GuHCl, D_m varies from 1.52 ± 0.01 to 1.62 ± 0.01 (Fig. 2, bottom), suggesting that, by these arguments, the chain is slightly more expanded than would be expected for a self-avoiding random walk polymer.

The equilibrium, ensemble unfolded state of protein L cannot be accessed without interference from a significant population of folded molecules below ~3.5 M GuHCl. This limitation restricts the precision with which equilibrium SAXS measurements can be used to determine the extent of unfolded state contraction. To broaden this range and better examine the extent to which unfolded protein L contracts as the denaturant concentration is reduced, we have employed nonequilibrium, time-resolved measurements to measure the dimensions of the unfolded state transiently populated at low denaturant prior to refolding. To perform such experiments, we have employed a conventional rapid-mixing apparatus that, using a variety of flow speeds and delay lines, provided access to sampling times as short as 4 ms. In our protocol, the unfolded protein in high denaturant concentration was rapidly diluted, and scattering was measured during and after the continuous-flow period. As the earliest time points collected are rapid relative to the protein's folding kinetics ($t_{fold} > 90$ ms under the conditions employed[31]), these measurements provide access to the unfolded state effectively without "contamination" due to scattering from native protein. This said, the transiently populated ensemble observed under these nonequilibrium conditions could differ from the unfolded state populated at equilibrium under the same conditions. Ensemble FRET measurements performed on protein L show, however, that a high transfer-efficiency state forms within a few microseconds.[17] This result indicates that unfolded protein L equilibrates to new solvent condition well within the 4-ms dead time of our experiments. These kinetic measurements thus provide an alternative strategy to smFRET for measuring properties of the unfolded state under conditions favoring the native state.

The unfolded state populated transiently prior to refolding at low denaturant concentrations appears indistinguishable by SAXS from the state observed at equilibrium under more highly denaturing conditions. Specifically, normalized scattering profiles collected within ~4 ms of rapid dilution to 0.67 and 1.3 M GuHCl, conditions well below the protein's denaturation midpoint, are effectively identical with those of the equilibrium unfolded state at 4 M GuHCl (Fig. 4, top). Consistent with this, the radii of gyration determined via Guinier analysis of these scattering profiles, 24.9 ± 1.12 Å and 23.5 ± 2.1 Å at 0.67 and 1.3 M GuHCl, respectively (Fig. 4, middle), are effectively indistinguishable from the 23.7 ± 0.4 observed at equilibrium under the 4-M GuHCl initial conditions. These values then decay to that of the native protein with folding time constants within error (albeit poorly constrained by the data) of those determined using fluorescence (Fig. S5). Analysis of the fractal dimensions of protein L under these conditions (Fig. 4, bottom) likewise suggests that its unfolded state remains expanded at low denaturant, with the D_m of the unfolded states observed at equilibrium at 4 M GuHCl and transiently at 1.3 and 0.67 M GuHCl all falling between 1.63 ± 0.07 and 1.55 ± 0.03, which is far below the $D_m = 3$ expected for a compact globule. Thus, at least as probed by SAXS, the dimensions and fractal dimensionality of unfolded protein L are independent of denaturant concentration to as low as 0.67 M GuHCl, an observation that contrasts sharply with the significant apparent contraction observed via smFRET.

Discussion

The SAXS profiles of unfolded protein L do not change appreciably as the concentration of GuHCl is reduced from 7.4 to 0.67 M, the lowest denaturant concentration we have investigated. This implies, in turn, that the shape and dimensions of the unfolded protein remain effectively fixed over this broad range of solvent conditions, and although this conclusion is consistent with that of prior SAXS studies of protein L,[22] the apparent lack of unfolded state collapse is, as noted above, significantly at odds with the interpretation of prior smFRET studies of this same protein.[8,9]

The results presented here speak to a significant discrepancy in the literature. That is, while a number of studies employing both FRET (e.g.,

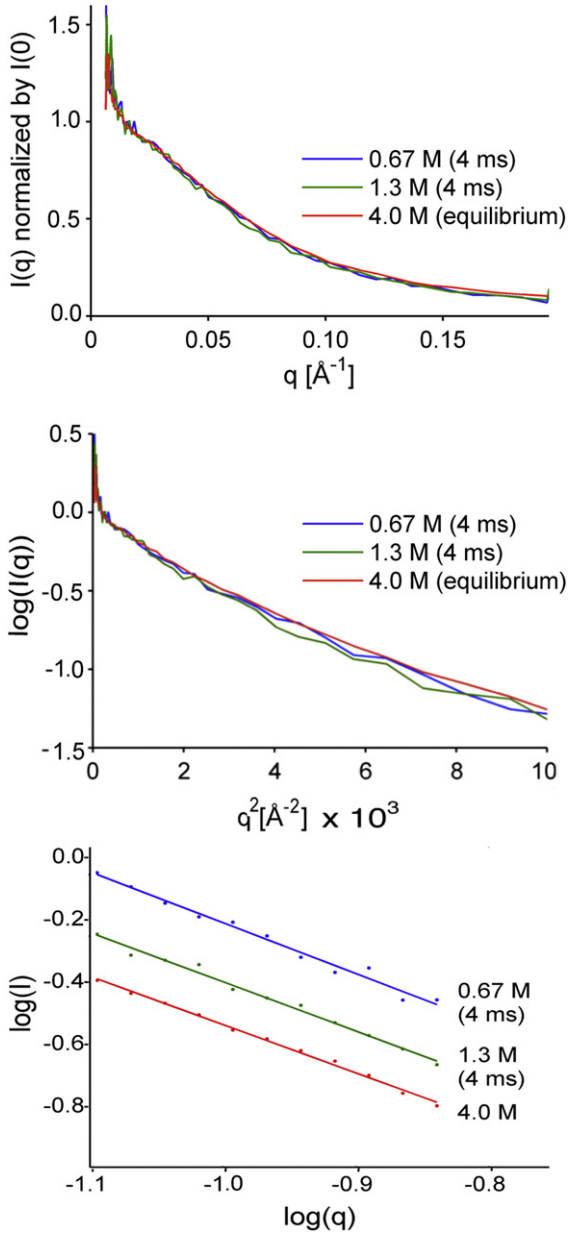

Fig. 4. The dimensions of the unfolded state of protein L transiently populated (prior to refolding) at low denaturant are, as determined by SAXS, effectively indistinguishable from those observed at equilibrium at higher denaturant. Shown are (top) raw scattering profiles [$I(0)$—normalized scattering intensity *versus* q] and (middle) Guinier plots for protein L at equilibrium at 4 M GuHCl and 4 ms after jumps from high denaturant to 0.67 or 1.3 M GuHCl. These data were acquired at the APS at 22 °C. (Bottom) Finally, analysis of the fractal dimensions of unfolded protein L likewise suggests that its unfolded state remains expanded at low denaturant as the slopes of log(I) *versus* log (q) plots (corresponding to the fractal dimension, D_m[29,30]) collected at 4 M GuHCl at equilibrium and transiently at 1.3 and 0.67 M GuHCl all fall between 1.63 ± 0.07 and 1.55 ± 0.03. The observed slopes approximate the value of 1.7 expected for an excluded volume random coil and fall far below the value of 3 expected for a compact globule.

Refs. 1,2,4–7,10–13,16,17, and 33) and fluorescence correlation spectroscopy[34,35] suggest that chain collapse at low denaturant is a near-universal property of unfolded polypeptides, a large body of earlier SAXS literature suggests such collapse is uncommon among proteins of under ~150 residues. For example, kinetic measurements performed on ubiquitin,[24] acylphosphatase,[24] superoxide dismutase (C. Kayatekin and O. Bilsel, personal communication), and barnase[27] suggest that, as is true for protein L, these proteins do not measurably contract during the burst phase of rapid-mixing experiments to low denaturant conditions. A similar lack of compaction is found in equilibrium SAXS studies of reduced, non-folding analogs of RNase A[26,30] and hen egg white lysozyme[21] at low pH; the dimensions of these analogs, which were produced by reduction of the two protein's disulfide bonds, remain at their high denaturant values to 0 and 0.3 M GuHCl, respectively, the lowest denaturant concentrations investigated. (At neutral pH, however, changes in the scattering curve of RNase A are observed, which may be consistent with a shift to more compact conformations at lower denaturant.[30]) Indeed, we are only aware of a handful of reported counterexamples among proteins of less than ~150 residues. One is a protein G variant that is reported to undergo unfolded state contraction under equilibrium conditions.[18] The wild-type protein, however, which differs from this construct by only a single residue, does not exhibit any significant evidence of collapse. Additional examples are disulfide-intact RNase and lysozyme, which are seen to collapse in time-resolved SAXS experiments.[19,36] These proteins, however, fold with complex, multi-state kinetics and thus may be distinct from the simple, two-state protein investigated here. Indeed, the time resolution of the relevant SAXS studies is poor relative to the faster refolding processes of each protein, suggesting that, for these proteins, the distinction between unfolded state contraction and the formation of an authentic, partially folded intermediate may be semantic.

In contrast to the situation at low denaturant, SAXS-based reports of collapse at zero denaturant are somewhat more common. These include studies of single-chain monellin (a 96-residue protein created by the fusion of two normally unlinked chains[20]), the 104-residue, heme-containing cytochrome *c*,[37,38] and the 153-residue apomyoglobin,[25] all of which contract in the dead time of rapid pH jumps under zero denaturant conditions. Likewise, the drkN SH3 domain, for which the dimensions of the unfolded state at zero denaturant were estimated using equilibrium SAXS measurements performed on a mixed population,[39] and the equilibrium-unfolded, truncated Δ131Δ variant of staphylococcal nuclease[40] are found to contract at zero denaturant, although both proteins expand to

random-coil dimensions at the lowest non-zero denaturant conditions that have been reported in the literature (2 M GuHCl and 8 M urea, respectively). Of note, however, the compaction of Δ131Δ is abolished by a single point mutant, producing dimensions that, in the absence of denaturant, are within error of those reported in 8 M urea.[40] Moreover, since this single point mutation shifts Δ131Δ from a compact denatured (or intermediate) state to an expanded, random-coil state (with no change in solvent conditions), it serves as an excellent control: for the low-denaturant unfolded states of this protein at least, SAXS easily distinguishes collapse from non-collapse, an observation that serves as an important control experiment for the studies described here. Finally, although no collapse is observed for ubiquitin at non-zero denaturant concentrations at room temperature,[24] a significant, rapid collapse phase is seen when the protein refolds at −20 °C in the presence of 45% ethylene glycol.[41]

We recognize that the suggestion that the unfolded state of protein L and, by extension, those of many other single-domain proteins, remains expanded to even quite low denaturant concentrations is controversial. In particular, a large number of theoretical and simulations-based studies suggest that chemically unfolded proteins undergo (nonspecific) collapse as the denaturant concentration is reduced (e.g., Refs. 9 and 42–44). More generally, the gradual equilibrium collapse suggested by, to the best of our knowledge, all prior smFRET studies is consistent with the physics of homopolymers, which are likely to collapse to form a nonspecific globule if the solvent quality falls sufficiently low (see detailed discussion in Ref. 28). We note, however, that the lack of a well-populated collapsed state suggested by SAXS is not inconsistent with physical principles. In particular, the lack of collapse may reflect important differences between the physics of proteins, which are, of course, sequence-specific heteropolymers containing both hydrophilic and hydrophobic elements in close concert, and the physics of simpler, more homogeneous polymers. To see this, consider the following. All compact states, including the native state, are stabilized by the burial of hydrophobic surface and destabilized by a loss of chain entropy and backbone desolvation. The native state, however, is also stabilized by interactions that require intricate and accurate spatial arrangement of the polypeptide chain, such as backbone hydrogen bonding and some specific side-chain interactions, which are likely incompletely formed in other collapsed states. If the thermodynamic balance is such that these interactions are required in order to compensate for the chain entropy loss and cost of backbone desolvation, these other collapsed states will not be stable, and the native state will be the only compact state that is stable relative to the expanded, unfolded ensemble. Thus, the lack of nonspecific collapse implied by the SAXS results can also have a physically reasonable interpretation.

What, then, is the origin of the discordant interpretations of so many SAXS and smFRET studies? A potential answer is that prior interpretations of SAXS scattering data and/or FRET efficiencies in terms of unfolded-state dimensions are subject to some hitherto unrecognized source (or sources) of systematic error. With regard to the SAXS results, however, we have failed to identify any experimental bias artifact that might account for this discrepancy. Moreover, such an artifact would have to exactly counterbalance a changing scattering profile such that the dimensions of the unfolded protein coincidentally appear unchanging to within experimental error. Indeed, the relationship between size, shape, and scattering involves well-understood physics and requires no obvious denaturant-dependent experimental variables other than the subtraction of background scattering, which is independently measured under each set of experimental conditions. Consistent with this, none of the several potential failings of SAXS previously suggested to account for the observed discrepancy appear to be valid. These include low contrast,[8] which, as shown here, does not significantly degrade our ability to measure unfolded-state dimensions, especially at the low denaturant conditions where the signal-to-noise ratio of SAXS is best and the SAXS–smFRET discrepancy is greatest; interparticle interference leading to an overestimation of R_g that perfectly cancels out the putative collapse,[9] which is ruled out by the differing protein concentrations employed here; and, finally, an inability to measure the dimensions of the unfolded state under strongly native conditions,[43] which is overcome using the nonequilibrium approaches applied here and previously (e.g., Refs. 22 and 24).

In contrast to the situation with SAXS, the interpretation of FRET efficiency in terms of molecular dimensions depends on several parameters that, at least in theory, are denaturant dependent and thus could lead to the appearance of a GuHCl-dependent change in dimensions when, in fact, there is none. The relevant control experiments have been performed, however, suggesting that these effects are minimal. For example, although the spectral overlap of the donor and acceptor, the index of refraction of the intervening medium, the quantum yield of the donor in the absence of acceptor, and various orientational averaging effects, all of which contribute to the Förster distance, are all at least weak functions of denaturant concentration, smFRET studies of native proteins and rigid polyproline constructs suggest that these effects are too small to account for the observed discrepancy (e.g., Refs. 8,9, and 45–47). Theory-

based analysis likewise suggests that the GuHCl dependence of conformational averaging in the flexible, highly dynamic unfolded state, which would be affected by denaturant-driven changes in solvent viscosity, is similarly too small to account for the observed change in transfer efficiency.[48] Finally, FRET spectroscopy is, in a sense, a perturbative experiment—the approach requires the addition of a donor/acceptor dye pair, which could, in theory, associate with themselves or the polypeptide chain due to interactions that are disrupted by denaturant. Arguing against this, however, is the observation that apparent collapse has been seen using several chemically distinct dye pairs (e.g., Refs. 4–7,12, and 33); indeed, in one example, the smFRET-derived appearance of collapse has been observed using two different dye pairs to study a single protein.[4,5] We are thus currently at a loss to explain the substantive disagreement between the SAXS results reported here and elsewhere and the results of prior smFRET studies.

While we have not identified the source of the SAXS–smFRET discrepancy, we nevertheless can draw several important conclusions from this and prior studies. First, there are many single-domain proteins for which the SAXS data appear strongly inconsistent with a denatured state that gradually collapses with decreasing denaturant concentration. Second, among the proteins that, by SAXS, do not appear to undergo gradual collapse as the denaturant concentration is reduced is protein L, an example for which the interpretation of its SAXS and smFRET characterizations unambiguously disagrees. The results presented here, which confirm earlier studies of this same protein,[22] thus argue that the disagreement between SAXS and smFRET is serious and unresolved and that suggestions that stable collapsed states are an important step in the rapid folding of proteins (e.g., Ref. 49) or that the large majority of polypeptide sequences undergo compaction at low denaturant (e.g., Refs. 16 and 35) may be premature. In short, we caution that our understanding of the physics of collapse of polypeptide sequences, particularly the factors that determine whether an unfolded protein will collapse or not, is far from being satisfactory and comprehensive.

Materials and Methods

Sample preparation

For measurements performed at the APS at the Argonne National Laboratories, protein L (pseudo-wild type—see Table S1) was expressed using *Escherichia coli* BL 21 cells and was purified either by anion exchange (Q Sepharose Fast Flow) or reverse-phase HPLC. For equilibrium measurements, sample purity was >95% with the minor product weighing an additional 64 Da according to analytical HPLC coupled to an electrospray time-of-flight mass spectrometry using an Agilent 1100 Series LC/MSD Trap. For the kinetic measurement, the protein purity was ~90% estimated from analytical HPLC followed by matrix-assisted laser desorption/ionization (PerSeptive Biosystems Voyager-DE Biospectrometry Workstation), with the minor products weighing an additional 20–62 Da. Protein concentration was determined using $\varepsilon_{280\ nm} = 9530\ M^{-1}\ cm^{-1}$.

For measurements performed at CHESS, we used a second protein L construct 72 residues in length containing a hexahistidine affinity tag (see Table S1). This protein was expressed and purified as previously described[50] with minor changes: lysis was performed by sonication, and purification was carried out using a nickel column (HiTrap Chelating HP 1 ml, GE Healthcare). Equilibrium fluorescence emission spectra for the purified protein L were measured in varying concentrations of guanidinium chloride buffered with 50 mM sodium phosphate, pH 7.0, with excitation at 280 nm. The denaturation curve was constructed from the shift of the emission peak, showing a very similar trend to that reported by Scalley *et al.*[51]

SAXS measurements

The equilibrium SAXS experiments performed at APS were conducted on the BioCAT beamline at the Advance Photon Source (APS) as previously described.[52] Each sample is loaded into a 1.5-mm capillary tube controlled by a Hamilton titrator with the sample holder thermostatically controlled at 22 °C. Buffer data are collected immediately prior to taking sample data under continuous-flow conditions to minimize radiation damage. For the series conducted from 1 to 7.4 M GuHCl, an automated 96-well protocol was used. Unless indicated, data were collected at 22 °C in 50 mM Hepes, pH 7.5, and 100 mM NaCl.

The time-resolved SAXS experiments were performed at the BioCAT beamline at APS using a Pilatus 100K photon counting detector (Dectris). Starting conditions were 13 mg/ml protein L in 4 M GuHCl. This sample was diluted with GuHCl-free buffer to 0.67 and 1.3 M GuHCl using a BioLogic brand SFM-400 stopped-flow mixer in continuous-flow mode. A 4-ms dead time was achieved by installing a microvolume mixer (MEC 22998) in the device and confirmed by time-resolved UV-absorption measurements of the reduction of 2,6-dichlorophenolindophenol by ascorbic acid.[53] Data points at longer times were obtained using delay loops. Separate measurements on the native and unfolded protein did not exhibit measurable radiation damage at 7-fold slower flow speeds.

The equilibrium SAXS measurements performed at CHESS covered a range of 0 M to 5.3 M GuHCl with protein concentration fixed at 10 mg/ml. Protein and matching buffer solutions consisted of 50 mM phosphate buffer, pH 7.0. Prior to taking data, the samples were centrifuged at 10,000 RPM for 10 min at 4 °C. SAXS data were collected at the G1 beamline at CHESS at 10.5 keV with a sample-to-detector distance of 1 m. Samples were placed in 1-mm-thick parylene-coated aluminum cells with silicon nitride windows to minimize background scattering. The sample holder was held at constant temperature (±0.2 °C) using a thermoelectric heat pump. Data were acquired at both 5 °C and 20 °C.

Data analysis

Radii of gyration were estimated from the raw scattering profiles using standard Guinier analysis.[54] Fractal dimensions (Figs. 2 and 4) were determined as outlined in the literature.[29,30] The chemical melts (plots of equilibrium R_g *versus* denaturant concentration; Fig. 3) were fit to standard, two-state linear free-energy relationships except that they were fitted to R_g^2 to capture the geometric averaging that occurs when determining the averaged R_g of mixed populations (see, e.g., Ref. 23).

Ensemble fluorescence measurements

The refolding of protein L (APS construct) was performed using the same apparatus as for the SAXS-monitored kinetic measurements but using a Biologic micro-cuvette in place of the X-ray capillary. Fluorescence of the tryptophan was measured using a 280- to 290-nm excitation wavelength and a 300- to 400-nm emission filter (Biologic MOS-200).

Acknowledgements

We wish to thank L. Guo, O. Bilsel, G. Gawlack, and members of S. Kent laboratory for providing prepublication results and expert assistance. This work was supported by research grants from the National Institutes of Health (NIH) [GM55694 (T.S.), GM080515 (G.H.), and EB002046 (K.W.P.)] and by an NIH training grant [GM-008267 (Cornell)]. Use of the APS was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract no. W-31-109-ENG-38. BioCAT is an NIH-supported Research Center RR-08630. CHESS is supported by the National Science Foundation and the NIH/National Institute of General Medical Sciences under National Science Foundation award DMR-0225180. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Center for Research Resources or the NIH.

Supplementary Data

Supplementary data to this article can be found online at doi:10.1016/j.jmb.2012.01.016

References

1. Kohn, J. E., Millett, I. S., Jacob, J., Zagrovic, B., Dillon, T., Cingel, N. *et al.* (2004). Do the dimensions of chemically unfolded proteins differ significantly from the expectations of a random-coil model? *Proc. Natl Acad. Sci. USA*, **101**, 12491–12496.
2. McCarney, E. R., Kohn, J. E. & Plaxco, K. W. (2005). Is there or isn't there? The case for (and against) residual structure in chemically denatured proteins. *Crit. Rev. Biochem. Mol. Biol.* **40**, 181–189.
3. Haran, G. (2011). How, when and why proteins collapse: the relation to folding. *Curr. Opin. Struct. Biol.* **22**, 14–20.
4. Deniz, A. A., Laurence, T. A., Beligere, G. S., Dahan, M., Martin, A. B., Chemla, D. S. *et al.* (2000). Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proc. Natl Acad. Sci. USA*, **97**, 5179–5184.
5. Laurence, T. A., Kong, X. X., Jäger, M. & Weiss, S. (2005). Probing structural heterogeneities and fluctuations of nucleic acids and denatured proteins. *Proc. Natl Acad. Sci. USA*, **102**, 17348–17353.
6. Schuler, B., Lipman, E. A. & Eaton, W. A. (2002). Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature*, **419**, 743–747.
7. Kuzmenkina, E. V., Heyes, C. D. & Nienhaus, G. U. (2006). Single-molecule FRET study of denaturant induced unfolding of RNase H. *J. Mol. Biol.* **357**, 313–324.
8. Sherman, E. & Haran, G. (2006). Coil–globule transition in the denatured state of a small protein. *Proc. Natl Acad. Sci. USA*, **103**, 11539–11543.
9. Merchant, K. A., Best, R. B., Louis, J. M., Gopich, I. V. & Eaton, W. A. (2007). Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations. *Proc. Natl Acad. Sci. USA*, **104**, 1528–1533.
10. Tezuka-Kawakami, T., Gell, C., Brockwell, D. J., Radford, S. E. & Smith, D. A. (2006). Urea-induced unfolding of the immunity protein Im9 monitored by spFRET. *Biophys. J.* **91**, L42–L44.
11. Hoffmann, A., Kane, A., Nettels, D., Hertzog, D. E., Baumgartel, P., Lengefeld, J. *et al.* (2007). Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy. *Proc. Natl Acad. Sci. USA*, **104**, 105–110.
12. Huang, F., Sato, S., Sharpe, T. D., Ying, L. M. & Fersht, A. R. (2007). Distinguishing between cooperative and unimodal downhill protein folding. *Proc. Natl Acad. Sci. USA*, **104**, 123–127.
13. Mukhopadhyay, S., Krishnan, R., Lemke, E. A., Lindquist, S. & Deniz, A. A. (2007). A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. *Proc. Natl Acad. Sci. USA*, **104**, 2649–2654.
14. Muller-Spath, S. M., Sorannol, A., Hirschfeld, V., Hofmann, H., Ruegger, S., Reymond, L. *et al.* (2010). Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl Acad. Sci. USA*, **107**, 14609–14614.
15. Ferreon, A. C. M., Moran, C. R., Gambin, Y. & Deniz, A. A. (2010). Single-molecule fluorescence studies of intrinsically disordered proteins. *Methods Enzymol.* **472**, 179–204.
16. Moglich, A., Joder, K. & Kiefhaber, T. (2006). End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation. *Proc. Natl Acad. Sci. USA*, **103**, 12394–12399.

17. Waldauer, S. A., Bakajin, O., Ball, T., Chen, Y., DeCamp, S. J., Kopka, M. *et al.* (2008). Ruggedness in the folding landscape of protein L. *HFSP J.* **2**, 388–395.
18. Smith, C. K., Bu, Z. M., Anderson, K. S., Sturtevant, J. M., Engelman, D. M. & Regan, L. (1996). Surface point mutations that significantly alter the structure and stability of a protein's denatured state. *Protein Sci.*, 2009–2019.
19. Kimura, T., Akiyama, S., Uzawa, T., Ishimori, K., Morishima, I., Fujisawa, T. & Takahashi, S. (2005). Specifically collapsed intermediate in the early stage of the folding of ribonuclease A. *J. Mol. Biol.* **350**, 349–362.
20. Kimura, T., Uzawa, T., Ishimori, K., Morishima, I., Takahashi, S., Konno, T. *et al.* (2005). Specific collapse followed by slow hydrogen-bond formation of β-sheet in the folding of single-chain monellin. *Proc. Natl Acad. Sci. USA*, **102**, 2748–2753.
21. Hoshino, M., Hagihara, Y., Hamada, D., Kataoka, M. & Goto, Y. (1997). Trifluoroethanol-induced conformational transition of hen egg-white lysozyme studied by small-angle X-ray scattering. *FEBS Lett.* **416**, 72–76.
22. Plaxco, K. W., Millett, I. S., Segel, D. J., Doniach, S. & Baker, D. (1999). Polypeptide chain collapse can occur concomitantly with the rate limiting step in protein folding. *Nat. Struct. Biol.* **6**, 554–557.
23. Millet, I. S., Townsley, L., Chiti, F., Doniach, S. & Plaxco, K. W. (2002). Equilibrium collapse and the kinetic foldability of proteins. *Biochemistry*, **41**, 321–325.
24. Jacob, J., Krantz, B., Dothager, R. S., Thiyagarajan, P. & Sosnick, T. R. (2004). Early collapse is not an obligate step in protein folding. *J. Mol. Biol.* **338**, 369–382.
25. Uzawa, T., Kimura, T., Ishimori, K., Morishima, I., Matsui, T., Ikeda-Saito, M. *et al.* (2006). Time-resolved small-angle X-ray scattering investigation of the folding dynamics of heme oxygenase: implication of the scaling relationship for the submillisecond intermediates of protein folding. *J. Mol. Biol.* **357**, 997–1008.
26. Jacob, J., Dothager, R. S., Thiyagarajan, P. & Sosnick, T. R. (2007). Fully reduced ribonuclease A does not expand at high denaturant concentration or temperature. *J. Mol. Biol.* **367**, 609–615.
27. Konuma, T., Kimura, T., Matsumoto, S., Goto, Y., Fujisawa, T., Fersht, A. R. & Takahashi, S. (2011). Time-resolved small angle X-ray scattering study of the folding dynamics of barnase. *J. Mol. Biol.* **405**, 1284–1294.
28. Ziv, G. & Haran, G. (2009). Protein folding, protein collapse, and Tanford's transfer model: lessons from single-molecule FRET. *J. Am. Chem. Soc.* **131**, 2942–2947.
29. Johansen, D., Trewhella, J. & Goldenberg, D. P. (2011). Fractal dimension of an intrinsically disordered protein: small-angle X-ray scattering and computational study of the bacteriophage λ N protein. *Protein Sci.* **20**, 1955–1970.
30. Wang, Y., Trewhella, J. & Goldenberg, D. P. (2008). Small-angle x-ray scattering of reduced ribonuclease A: effects of solution conditions and comparisons with a computational model of unfolded proteins. *J. Mol. Biol.* **377**, 1576–1592.
31. Scalley, M. L. & Baker, D. (1997). Protein folding kinetics exhibit an Arrhenius temperature dependence when corrected for the temperature dependence of protein stability. *Proc. Natl Acad. Sci. USA*, **94**, 10636–10640.
32. Millett, I. S., Doniach, S. & Plaxco, K. W. (2002). Towards a taxonomy of the denatured state: small angle scattering studies of unfolded proteins. *Adv. Protein Chem.* **62**, 241–262.
33. Sinha, K. K. & Udgaonkar, J. B. (2005). Dependence of the size of the initially collapsed form during the refolding of barstar on denaturant concentration: evidence for a continuous transition. *J. Mol. Biol.* **353**, 704–718.
34. Crick, S. L., Jayaraman, M., Frieden, C., Wetzel, R. & Pappu, R. V. (2006). Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. *Proc. Natl Acad. Sci. USA*, **103**, 16764–16769.
35. Teufel, D. P., Johnson, C. M., Lum, J. K. & Neuweiler, H. (2011). Backbone-driven collapse in unfolded protein chains. *J. Mol. Biol.* **409**, 250–262.
36. Segel, D. J., Bachmann, A., Hofrichter, J., Hodgson, K. O., Doniach, S. & Kiefhaber, T. (1999). Characterization of transient intermediates in lysozyme folding with time-resolved small-angle X-ray scattering. *J. Mol. Biol.* **288**, 489–499.
37. Pollack, L., Tate, M. W., Darnton, N. C., Knight, J. B., Gruner, S. M., Eaton, W. A. & Austin, R. H. (1999). Compactness of the denatured state of a fast-folding protein measured by submillisecond small-angle x-ray scattering. *Proc. Natl Acad. Sci. USA*, **96**, 10115–10117.
38. Akiyama, S., Takahashi, S., Kimura, T., Ishimori, K., Morishima, I., Nishikawa, Y. & Fujisawa, T. (2002). Conformational landscape of cytochrome *c* folding studied by microsecond-resolved small angle x-ray scattering. *Proc. Natl Acad. Sci. USA*, **99**, 1329–1334.
39. Choy, W. Y., Mulder, F. A. A., Crowhurst, K. A., Muhandiram, D. R., Millett, I. S., Doniach, S. *et al.* (2002). Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *J. Mol. Biol.* **31**, 101–112.
40. Flanagan, J. M., Kataoka, M., Fujisawa, T. & Engelman, D. M. (1993). Mutations can cause large changes in the conformation of a denatured protein. *Biochemistry*, **32**, 10359–10370.
41. Larios, E., Li, J. S., Schulten, K., Kihara, H. & Gruebele, M. (2004). Multiple probes reveal a native-like intermediate during low-temperature refolding of ubiquitin. *J. Mol. Biol.* **340**, 115–125.
42. Canchi, D. R., Paschek, D. & Garcia, A. E. (2010). Equilibrium study of protein denaturation by urea. *J. Am. Chem. Soc.* **132**, 2338–2344.
43. O'Brien, E. P., Brooks, B. R. & Thirumalai, D. (2009). Molecular origin of constant *m*-values, denatured state collapse, and residue-dependent transition midpoints in globular proteins. *Biochemistry*, **48**, 3743–3754.
44. Pace, N. C., Huyghues-Despointes, B. M., Fu, H., Takano, K., Scholtz, J. M. & Grimsley, G. R. (2010). Urea denatured state ensembles contain extensive secondary structure that is increased in hydrophobic proteins. *Protein Sci.* **19**, 929–943.

45. Kuzmenkina, E. V., Heyes, C. D. & Nienhaus, G. U. (2005). Single-molecule Forster resonance energy transfer study of protein dynamics under denaturing conditions. *Proc. Natl Acad. Sci. USA*, **102**, 15471–15476.
46. Schuler, B., Lipman, E. A., Steinbach, P. J., Kumke, M. & Eaton, W. A. (2008). Polyproline and the spectroscopic ruler revisited with single-molecule fluorescence. *Proc. Natl Acad. Sci. USA*, **102**, 2754–2759.
47. Jha, S. K., Dhar, D., Krishnamoorthyc, G. & Udgaonkara, J. B. (2009). Continuous dissolution of structure during the unfolding of a small protein. *Proc. Natl Acad. Sci. USA*, **106**, 11113–11118.
48. Makarov, D. E. & Plaxco, K. W. (2009). Measuring distances within unfolded biopolymers using fluorescence resonance energy transfer: the effect of polymer chain dynamics on the observed fluorescence resonance energy transfer efficiency. *J. Chem. Phys.* **131**; Art No 085105.
49. Sinha, K. K. & Udgaonkar, J. B. (2009). Early events in protein folding. *Curr. Sci.* **96**, 1053–1070.
50. Gu, H., Yi, Q. A., Bray, S. T., Riddle, D. S., Shiau, A. K. & Baker, D. (1995). A phage display system for studying the sequence determinants of protein folding. *Protein Sci.* **4**, 1108–1117.
51. Scalley, M. L., Yi, Q., Gu, H. D., McCormack, A., Yates, J. R. & Baker, D. (1997). Kinetics of folding of the IgG binding domain of peptostreptococcal protein L. *Biochemistry*, **36**, 3373–3382.
52. Pan, T. & Sosnick, T. R. (2005). Structural analysis of RNA and RNA–protein complexes by small angle X-ray scattering. In *Handbook of RNA Biochemistry* (Hartmann, R. K., Bindereif, A., Schon, A. & Westhof, E., eds), pp. 385–397, Wiley-VCH Verlag, Weinheim, Germany.
53. Roh, J. H., Guo, L., Kilburn, J. D., Briber, R. M., Irving, T. & Woodson, S. A. (2010). Multistage collapse of a bacterial ribozyme observed by time-resolved small-angle X-ray scattering. *J. Am. Chem. Soc.* **132**, 10148–10154.
54. Svergun, D. I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503.