Published in final edited form as:

Curr Opin Struct Biol. 2012 February ; 22(1): 14–20. doi:10.1016/j.sbi.2011.10.005.

How, when and why proteins collapse: the relation to folding

Gilad Haran

Chemical Physics Department, Weizmann Institute of Science, Rehovot 76100, Israel,
gilad.haran@weizmann.ac.il

Abstract

Unfolded proteins under strongly-denaturing conditions are highly expanded. However, when the conditions are more close to native, an unfolded protein may collapse to a compact globular structure distinct from the folded state. This transition is akin to the coil-globule transition of homopolymers. Single-molecule FRET experiments have been particularly conducive in revealing the collapsed state under conditions of coexistence with the folded state. The collapse can be even more readily observed in natively unfolded proteins. Time-resolved studies, using FRET and small-angle scattering, have shown that the collapse transition is a very fast event, probably occurring on the sub-microsecond time scale. The forces driving collapse are likely to involve both hydrophobic and backbone interactions. The loss of configurational entropy during collapse makes the unfolded state less stable compared to the folded state, thus facilitating folding.

Introduction

Unfolded proteins are often depicted as highly expanded random coils. These structures are by necessity the starting points for the folding reaction, and the pathways they take on the way to the folded state motivate this review. Since a folded (globular) protein is compact, its folding reaction must involve a transition from an expanded structure to a collapsed one. In principle, this so-called collapse transition may occur concomitantly with the folding transition. However, much evidence, both from experiment and from theory, suggests a collapse transition that is distinct from folding (Figure 1).

Polymer collapse, also termed the coil-globule transition, is a well-known phenomenon in the physics of chain molecules. The collapse of homopolymers has been discussed extensively in classical textbooks. The size of polymers is governed by universal scaling laws [1], with statistics that depend on the environment. Thus, in a good solvent the interaction between monomers is effectively repulsive, and a polymer explores an expanded ensemble of random coil states, with the radius of gyration, R_g , scaling like $N^{3/5}$, where N is the number of monomers. In contrast, in a poor solvent the effective monomer-monomer interaction is attractive, and a polymer is globular, so that its R_g scales like $N^{1/3}$. A continuous, second-order transition connects these two phases [2].

Proteins, which are heteropolymers, have been shown to obey these two limits. The size of highly denatured proteins indeed scales like $N^{3/5}$, while folded proteins typically scale as globules [3–5]. But there is also a globular state of proteins that is distinct from the folded state, and is more akin to the collapsed state of polymers, as it is predominantly disordered.

© 2011 Elsevier Ltd. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

This globular state is also relevant to intrinsically disordered proteins (IDPs), where it can be studied in the absence of the ‘perturbation’ due to the folding transition. In this review we will discuss the mounting evidence for the formation of a collapsed state that may coexist with the folded state under certain conditions, and the possible energetic role of the coil-globule transition in the folding reaction. We will also address the dynamics of the process, and finally ask how universal it is.

Equilibrium studies of collapse: folding proteins

Under certain equilibrium conditions (e.g. low concentration of a chemical denaturant) the collapsed state of a protein may be populated, typically coexisting with the folded state. This coexistence complicates attempts to observe collapsed proteins. As will be seen, single-molecule fluorescence spectroscopy detects distinct signals from folded and collapsed protein molecules, even when the two populations co-exist. In addition, due to the continuous nature of the collapse transition, the signature of this process may sometime also be detectable under conditions where the folded state is not populated, such as at intermediate levels of denaturants. Indeed, some ensemble fluorescence resonance energy transfer (FRET) experiments have been able to observe structural changes indicative of protein chain collapse in folding proteins [6–8]. The equilibrium collapse was revealed by dynamic light scattering [9], and in some small-angle x-ray scattering (SAXS) experiments (but see more on SAXS below) [10,11].

Single-molecule FRET experiments on freely-diffusing molecules offer a very convenient method to separate the contributions of the folded and denatured states at low denaturant concentrations and hence observe the collapsed state. Such experiments have offered the most consistent evidence for the collapse transition under equilibrium conditions [12–21]. In single-molecule FRET experiments [22], a laser beam is focused within a dilute solution of protein molecules. Each molecule is labeled with two fluorescent probes, a donor and an acceptor. The passage of a molecule through the laser beam leads to excitation of the donor probe, which may then either emit a photon or transfer the energy to the acceptor. By gauging the number of photons emitted from the donor and acceptor one can infer the FRET efficiency for each protein molecule [23]. A histogram is then constructed from the FRET efficiency values of hundreds or thousands of molecules. For a folded or a fully-denatured protein, a single peak should appear in this histogram at a high or low FRET efficiency, respectively. Under intermediate conditions that allow folded and unfolded molecules to coexist, two peaks show up in the FRET efficiency histogram. The position of the denatured-state peak typically shifts continuously to higher FRET efficiency values as solvent conditions are modified, which is a clear indication that the denatured protein molecules become more compact, or in other words undergo a collapse transition. While most single-molecule FRET experiments used chemical denaturants to tune the solvent quality, Schuler and coworkers showed that temperature can also drive the process [19], as in the classical experiments on polymers [24], although in this case it is the increase in temperature that leads to a more compact structure.

Computer simulations of chemical denaturation of folding proteins also consistently observe that at lower denaturant concentration chains are more compact than the fully expanded, random-coil like chains at high denaturant concentration [16,25–27]. O’Brien et al. used an implicit denaturant model to study equilibrium denaturation of two single-domain proteins that were examined by single-molecule FRET, and found very good agreement of the denaturant-dependent size of the protein between the experimental and simulated results [25,26]. Garcia and coworkers offered a detailed study of the unfolding of the miniprotein Trp-cage in explicit urea solutions, and found that greater urea concentration favored conformations with greater solvent exposure [27].

Equilibrium studies of collapse: intrinsically disordered proteins

Many proteins “defy” the common paradigm in structural biology and do not fold to a well-defined three-dimensional structure under native conditions. These are the so-called IDPs [28]. Some of these proteins do fold upon binding to a target [29], while others do not seem to fold under any known conditions. Much has been written in recent years about the connection between the folding behavior of IDPs and their activity. Further, it was recognized that proteins belonging to this group are in general characterized by low hydrophobicity and high charge density [3]. IDPs are ideal targets for studies of the collapse transition of unfolded proteins, as they make the separation of folded and unfolded data unnecessary. Deniz and coworkers [30] appreciated this advantage and studied a natively unfolded segment of the yeast protein Sup35. Under native condition this protein was significantly more compact than measured in high denaturant concentration. Nevertheless, fast chain dynamics were detected even within this compact state.

The role of charge in determining the size of IDPs was investigated by Pappu and coworkers, who used fluorescence correlation spectroscopy (FCS) to look at the dimensions of a family of arginine-rich IDPs, the protamines [31], under native conditions. The net charge per residue was found to be a good predictor of the size, with a change in R_g by as much as a factor of two between the least and most highly charged variants. Back to single-molecule FRET, Müller-Spätth et al. [32] studied the correlation between charge density and overall dimensions of two IDPs, comparing the results to those obtained with a folding protein. All three proteins gradually collapse when the denaturant concentration is lowered. Surprisingly, the authors find that in the case of the two IDPs, R_g grows again as concentration of the ionic denaturant guanidinium chloride is lowered below 1 M. The authors concluded that this is due to “release” of the proteins from electrostatic screening, as the extent of the observed expansion in buffer correlates well with the mean net charge on each chain, just as in the work of Mao et al. [31].

IDPs have also been used to study the effect of crowding on chain dimensions [33]. In this small-angle neutron scattering experiment, a clever contrast variation was used to enable studying size variations of the N protein of bacteriophage λ in the presence of a background of a large concentration of crowding proteins. Interestingly, analysis of the scattering function at a relatively low crowder concentration indicated compaction of the protein chain, as expected based on a computational model. Less obvious effects were seen at higher crowder concentrations, suggesting a polymer melt-like behavior of the solution, which requires further studies.

Collapse kinetics

The collapse transition is predicted to temporally precede folding [34]. Therefore, there have been many experimental attempts to measure the collapse time in proteins by rapidly changing the solution conditions from denaturing to native-like. Many of these experiments detect the changes of a local probe, such as tryptophan fluorescence, which only indirectly reports on the overall change in the size of the protein. Thus, as in previous sections, we focus here on experiments that can probe the magnitude of the structural change involved in protein collapse. A number of time-resolved SAXS experiments indicated significant chain compaction following either a pH jump or a denaturant concentration jump within a fast mixing device (Figure 2) [11,35–42]. In all of these experiments the collapse was apparently too fast to be directly resolved, even when using state-of-the-art continuous-flow instruments with a dead time of hundreds of microseconds [43]. Takahashi and coworkers [41] analyzed a series of these results, and found that the scaling of the R_g of the proteins with length at the

earliest measured point following mixing matched the expectation for a globular state, with an exponent of 0.35.

Attempts to measure the collapse time have also been made using time-resolved FRET experiments, in which the change in the FRET efficiency of a labeled protein was monitored as a function of time after a laser-induced temperature jump [44] or a fast mixing step [45–53]. An early temperature jump experiment resolved a collapse time of 100 μ s for cytochrome c [44]. Other experiments failed to directly resolve the fast transient due to collapse, and were only able to obtain an upper limit for the collapse time. Interestingly, a recent mixing experiment with a dead time as short as 4 μ s could not resolve the time scale of collapse of protein L [52]. The conclusion from all these experiments is that proteins collapse may be extremely fast, probably occurring in some cases on a sub-microsecond time scale. The experimental means to directly measure the structural changes accompanying collapse in real time require further development.

Relation of protein collapse and folding

So why do proteins collapse? The folding transition is at least partially driven by the need to bury hydrophobic amino acids within the interior of the protein. It is likely that the collapse transition in a poor solvent is driven by a similar requirement, which imposes specific structural constraints on the collapsed state. This picture was emphasized in early simulations of folding, which represented proteins in terms of heteropolymers with two types of monomers, polar and hydrophobic, with an effective attraction between the latter [34].

Recent theoretical and experimental work proposed a different, perhaps complementary, view of the forces responsible for collapse, which highlights the importance of the protein backbone rather than the side chains. Bolen and coworkers showed that the free energy cost to transfer a protein from denaturing to native conditions is dominated by the backbone contribution [9,54,55]. Simulations of Pappu and coworkers [56] on polyglycine and glycine-serine peptides, whose collapsed configurations under native conditions were assigned to backbone interactions, indicated that water is a poor solvent for peptide backbones. The collapsed configuration of glycine-serine peptides was directly demonstrated in time-resolved FRET experiments [57] and more recent FCS studies [58]. Particularly revealing was the comparison by Neuweiler and coworkers between the size of a polyglycine peptide in water to the size of same peptide when all backbone amide hydrogens were replaced with methyl groups [58]. It was found that the latter is much more expanded than the former, strongly pointing to the participation of backbone hydrogen bonds in the collapse transition. More on the forces involved in protein folding and collapse can be found in recent reviews [55,59].

The debate on the relative importance of side chains vs. backbone in protein collapse notwithstanding, the detailed equilibrium view of the collapse transition offered by single-molecule FRET experiments paves the way to a quantitative assessment of the energetic connection between collapse and folding. Ziv and Haran [60] proposed a method for the analysis of single-molecule FRET experiments, based on the theory of the coil-globule transition of polymers [2,61,62]. This analysis allows estimating the enthalpic and entropic contributions to the coil-globule transition. The enthalpic contribution depends on denaturant concentration. It may be attractive, favoring chain collapse, or repulsive, favoring expansion. The configurational entropy term arises from excluded volume interactions, and always favors expansion. Thus the balance between the two terms determines the state of the chain.

Ziv and Haran [60] analyzed twelve single-molecule FRET data sets measured on five different proteins [13–18]. For each of these proteins, they calculated the effect of denaturant on the “collapse free energy”, $\Delta G_{U \rightarrow C}$, which is the difference free energy for the transition from an unfolded state, U , to the putative maximally collapsed state, C . Figure 3 shows an example of the function $\Delta G_{U \rightarrow C}$, calculated from experimental results obtained with protein L [14]. $\Delta G_{U \rightarrow C}$ turned out to be essentially linear with denaturant concentration for all proteins studied. Surprisingly, the slope of this linear function was found in all cases to be very similar to the slope of the folding free energy as a function of denaturant concentration, which is the famous m -value [63].

This energetic connection between collapse and folding has a far-reaching implication for the propensity of a protein to fold. Chain expansion stabilizes the denatured state by increasing its configurational entropy, and is therefore responsible for the increased propensity of a protein to be unfolded in high denaturant concentration. Conversely, the collapsed state facilitates folding, as its configurational entropy is lower, and hence it is less stable with respect to the folded state (Figure 1). Many IDPs are polar and charged molecules, and they cannot spontaneously collapse even in the absence of denaturants. This property keeps them disordered until they are forced to fold, e.g. by binding to a partner protein.

A dissenting view from SAXS experiments

So far we have seen much evidence, mostly from FRET studies, that there is a coil-globule transition in the unfolded state of folding proteins, as well as in IDPs, when solution conditions are modulated. In clear contradiction to this picture, however, many equilibrium SAXS denaturation experiments find no evidence for changes in chain dimensions beyond the unfolding transition [38,64–68]. In addition, some time-resolved stopped-flow SAXS experiments have also failed to observe chain collapse following rapid change of conditions from denaturing to native [64,67,68]. The discrepancy is particularly startling in the case of protein L, which has been studied by both single-molecule FRET [14,16] and SAXS [64]1, with conflicting results. It is possible, though, that the solution lies in the difference in sampling of collapsed and expanded conformers by the two methods: FRET tends to weigh more compact conformations, while SAXS is biased by extended conformations [69]. It is hoped that solving this conundrum will teach us something new and fundamental about the unfolded state.

Conclusion

What remains to be studied in relation to protein collapse? A crucial issue, which we didn't discuss at all in this review, is the formation of secondary structure concomitantly with chain collapse [17,37–39,70]. Secondary structure may form as a by-product of collapse, or it may be its driving force. In this context, it has been suggested [71] (and corroborated by experiment [37]) that β -sheet proteins are more likely to experience an initially structure-less collapse. A second, and related question, is the thermodynamic nature of the coil-globule transition of proteins. While single-molecule FRET experiments suggest a continuous, second-order transition, as in homopolymers, it is possible that in some proteins this is in fact a first order transition, with an energy barrier separating the expanded and collapsed configurations [44,72]. A collapsed state separated by an energy barrier can be seen as a stable folding intermediate. However, a first-order collapse transition may arise even in homopolymers, e.g. when dewetting of the chain (i.e. formation of a water-less region

¹The results of this SAXS experiment have been reproduced recently (Yoo, T. Y., Meisberger, S., Hinshaw, J., Pollack, L., Haran, G., Sosnick, T. R., Plaxco, K., submitted).

around it) is the driving force [73]. Therefore, knowledge of the order of the transition might not be enough to distinguish a non-specific collapse from a 'specific' folding intermediate. New experimental methods, perhaps on the single-molecule level, will probably shed light on these open questions in the future. And since many of the experiments performed so far were done on relatively small, single-domain proteins, it will also be interesting to perform experiments on larger, multi-domain proteins, and find out whether the current concepts can be extended to these proteins as well.

A fruitful dialog between protein science and polymer physics has been going on for many years. We focused here on one important feature of polymers, the coil-globule transition, which has been found to be a useful concept for the description of conformational properties of unfolded proteins. Science sometimes works in circles. Thus, the interest in protein collapse has waxed and waned several times over the years. The appearance of ultrasensitive experiments, such as single-molecule FRET spectroscopy and FCS, and the discovery of IDPs have revived the interest in the collapse of protein chains under native or close-to-native conditions. Together with additional open questions in the protein folding field [74], there is no doubt that the collapse transition will keep us busy in the years to come.

Highlights

- Unfolded proteins are expanded under strongly denaturing conditions.
- They collapse when conditions are close-to-native.
- Single-molecule fluorescence shows that collapsed and folded states coexist.
- Time-resolved experiments reveal that collapse usually occurs much before folding.
- Both hydrophobic interactions and hydrogen bonding may be involved in driving collapse.

Acknowledgments

This work was supported by grant no. R01GM080515 from the NIH. G.H. thanks George Rose for many illuminating discussions, and Inbal Riven for help with graphics and useful comments.

References

1. de Gennes, P-G. Scaling concepts in polymer physics. Cornell university press: Ithaca; 1979.
2. Grosberg, AY.; Khokhlov, AR. Statistical Physics of Macromolecules. AIP Press: New York; 1994.
3. Uversky VN. Natively unfolded proteins: A point where biology waits for physics. *Protein Science*. 2002; 11:739–756. [PubMed: 11910019]
4. Wilkins DK, Grimshaw SB, Receveur V, Dobson CM, Jones JA, Smith LJ. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry*. 1999; 38:16424–16431. [PubMed: 10600103]
5. Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, Dothager RS, Seifert S, Thiyagarajan P, Sosnick TR, et al. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101:12491–12496. [PubMed: 15314214]
6. Lakshminanth GS, Sridevi K, Krishnamoorthy G, Udgaonkar JB. Structure is lost incrementally during the unfolding of barstar. *Nat Struct Biol*. 2001; 8:799–804. [PubMed: 11524685]
7. Lee JC, Engman KC, Tezcan FA, Gray HB, Winkler JR. Structural features of cytochrome c' folding intermediates revealed by fluorescence energy-transfer kinetics. *Proc Natl Acad Sci U S A*. 2002; 99:14778–14782. [PubMed: 12407175]

8. Huang F, Lerner E, Sato S, Amir D, Haas E, Fersht AR. Time-resolved fluorescence resonance energy transfer study shows a compact denatured state of the B domain of protein A. *Biochemistry*. 2009; 48:3468–3476. [PubMed: 19222162]
9. Holthauzen LM, Rosgen J, Bolen DW. Hydrogen bonding progressively strengthens upon transfer of the protein urea-denatured state to water and protecting osmolytes. *Biochemistry*. 2010; 49:1310–1318. [PubMed: 20073511] Dynamic light scattering is used to show collapse in water and osmolyte solutions. Analysis of the energetics involved in the process shows that the largest contribution is due to backbone-backbone, or hydrogen-bonding interactions.
10. Smith CK, Bu Z, Anderson KS, Sturtevant JM, Engelman DM, Regan L. Surface point mutations that significantly alter the structure and stability of a protein's denatured state. *Protein Sci*. 1996; 5:2009–2019. [PubMed: 8897601]
11. Arai M, Kondrashkina E, Kayatekin C, Matthews CR, Iwakura M, Bilsel O. Microsecond hydrophobic collapse in the folding of *Escherichia coli* dihydrofolate reductase, an alpha/beta-type protein. *J Mol Biol*. 2007; 368:219–229. [PubMed: 17331539] A SAXS study that identifies the collapse transition of the protein dihydrofolate reductase both in both equilibrium and time-resolved experiments. The collapse is too fast to resolve with the 300 μ s time resolution of this experiment.
12. Laurence TA, Kong X, Jager M, Weiss S. Probing structural heterogeneities and fluctuations of nucleic acids and denatured proteins. *Proc Natl Acad Sci U S A*. 2005; 102:17348–17353. [PubMed: 16287971]
13. Kuzmenkina EV, Heyes CD, Nienhaus GU. Single-molecule FRET study of denaturant induced unfolding of RNase H. *J Mol Biol*. 2006; 357:313–324. [PubMed: 16426636]
14. Sherman E, Haran G. Coil-globule transition in the denatured state of a small protein. *Proc Natl Acad Sci U S A*. 2006; 103:11539–11543. [PubMed: 16857738]
15. Tezuka-Kawakami T, Gell C, Brockwell DJ, Radford SE, Smith DA. Urea-induced unfolding of the immunity protein Im9 monitored by spFRET. *Biophys J*. 2006; 91:L42–L44. [PubMed: 16798813]
16. Merchant KA, Best RB, Louis JM, Gopich IV, Eaton WA. Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations. *Proc Natl Acad Sci U S A*. 2007; 104:1528–1533. [PubMed: 17251351]
17. Hoffmann A, Kane A, Nettels D, Hertzog DE, Baumgartel P, Lengefeld J, Reichardt G, Horsley DA, 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; 104:105–110. [PubMed: 17185422]
18. Hofmann H, Golbik RP, Ott M, Hubner CG, Ulbrich-Hofmann R. Coulomb forces control the density of the collapsed unfolded state of barstar. *Journal of Molecular Biology*. 2008; 376:597–605. [PubMed: 18164723]
19. Nettels D, Muller-Spath S, Kuster F, Hofmann H, Haenni D, Ruegger S, Reymond L, Hoffmann A, Kubelka J, Heinz B, et al. Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins. *Proc Natl Acad Sci U S A*. 2009 While most single-molecule FRET experiments use a chemical denaturant to change the state of an unfolded protein from collapsed to expanded, this study shows that temperature can be used as well. But surprisingly, it is an increase in temperature that leads to chain collapse.
20. Liu P, Meng X, Qu P, Zhao XS, Wang CC. Subdomain-specific collapse of denatured staphylococcal nuclease revealed by single molecule fluorescence resonance energy transfer measurements. *J Phys Chem B*. 2009; 113:12030–12036. [PubMed: 19678648]
21. Rieger R, Kobitski A, Sielaff H, Nienhaus GU. Evidence of a folding intermediate in RNase H from single-molecule FRET experiments. *Chemphyschem*. 2011; 12:627–633. [PubMed: 21344597]
22. Chen H, Rhoades E. Fluorescence characterization of denatured proteins. *Curr Opin Struct Biol*. 2008; 18:516–524. [PubMed: 18675353] This review provides an excellent introduction to the single-molecule fluorescence methods used to study protein folding, particularly in relation to structure and dynamics of the unfolded state.
23. Michalet X, Weiss S, Jager M. Single-molecule fluorescence studies of protein folding and conformational dynamics. *Chem Rev*. 2006; 106:1785–1813. [PubMed: 16683755]

24. Sun ST, Nishio I, Swislow G, Tanaka T. The Coil-Globule Transition - Radius of Gyration of Polystyrene in Cyclohexane. *Journal of Chemical Physics*. 1980; 73:5971–5975.
25. O'Brien EP, Ziv G, Haran G, Brooks BR, Thirumalai D. Effects of denaturants and osmolytes on proteins are accurately predicted by the molecular transfer model. *Proc Natl Acad Sci U S A*. 2008; 105:13403–13408. [PubMed: 18757747] This paper introduces an implicit method to include chemical denaturants in molecular dynamics simulations, using Tanford's transfer model to compute the relative contribution of each denatured-state conformer of a protein to the ensemble.
26. O'Brien EP, Brooks BR, Thirumalai D. Molecular origin of constant m-values, denatured state collapse, and residue-dependent transition midpoints in globular proteins. *Biochemistry*. 2009; 48:3743–3754. [PubMed: 19278261]
27. Canchi DR, Paschek D, Garcia AE. Equilibrium study of protein denaturation by urea. *J Am Chem Soc*. 2010; 132:2338–2344. [PubMed: 20121105] A thorough simulational study of the denaturation of a small protein. The contribution of van der Waals interactions between urea and the protein is found to dominate the free energy of unfolding.
28. Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu Rev Biophys*. 2008; 37:215–246. [PubMed: 18573080]
29. Wright PE, Dyson HJ. Linking folding and binding. *Curr Opin Struct Biol*. 2009; 19:31–38. [PubMed: 19157855]
30. Mukhopadhyay S, Krishnan R, Lemke EA, Lindquist S, Deniz AA. A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. *Proc Natl Acad Sci U S A*. 2007; 104:2649–2654. [PubMed: 17299036]
31. Mao AH, Crick SL, Vitalis A, Chicoine CL, Pappu RV. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc Natl Acad Sci U S A*. 2010; 107:8183–8188. [PubMed: 20404210] A series of natively-unfolded proteins with a varying charge density is studied. The lower the charge density, the more collapsed is the protein, as shown by both experiment and simulation.
32. Müller-Spätth S, Soranno A, Hirschefeld V, Hofmann H, Rüggera S, Reymonda L, Nettels D, Schuler B. Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc Natl Acad Sci U S A*. 2010; 107:14609–14614. [PubMed: 20639465]
33. Johansen D, Jeffries CM, Hammouda B, Trehwella J, Goldenberg DP. Effects of macromolecular crowding on an intrinsically disordered protein characterized by small-angle neutron scattering with contrast matching. *Biophys J*. 2011; 100:1120–1128. [PubMed: 21320458]
34. Camacho CJ, Thirumalai D. Kinetics and thermodynamics of folding in model proteins. *Proc Natl Acad Sci U S A*. 1993; 90:6369–6372. [PubMed: 8327519]
35. Pollack L, Tate MW, Finnefrock AC, Kalidas C, Trotter S, Darnton NC, Lurio L, Austin RH, Batt CA, Gruner SM, et al. Time resolved collapse of a folding protein observed with small angle x-ray scattering. *Phys Rev Lett*. 2001; 86:4962–4965. [PubMed: 11384392]
36. Pollack L, Tate MW, Darnton NC, Knight JB, Gruner SM, Eaton WA, Austin RH. Compactness of the denatured state of a fast-folding protein measured by submillisecond small-angle x-ray scattering. *Proc Natl Acad Sci U S A*. 1999; 96:10115–10117. [PubMed: 10468571]
37. 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 beta-sheet in the folding of single-chain monellin. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102:2748–2753. [PubMed: 15710881]
38. Kimura T, Akiyama S, Uzawa T, Ishimori K, Morishima I, Fujisawa T, Takahashi S. Specifically collapsed intermediate in the early stage of the folding of ribonuclease A. *J Mol Biol*. 2005; 350:349–362. [PubMed: 15935376]
39. 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 alpha-helical content and compactness. *Proc Natl Acad Sci U S A*. 2004; 101:1171–1176. [PubMed: 14711991]

40. Akiyama S, Takahashi S, Kimura T, Ishimori K, Morishima I, Nishikawa Y, Fujisawa T. Conformational landscape of cytochrome c folding studied by microsecond-resolved small-angle x-ray scattering. *Proc. Natl. Acad. Sci. USA*. 2002; 99:1329–1334. [PubMed: 11773620]
41. Uzawa T, Kimura T, Ishimori K, Morishima I, Matsui T, Ikeda-Saito M, Takahashi S, Akiyama S, Fujisawa T. 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*. 2006; 357:997–1008. [PubMed: 16460755]
42. Wu Y, Kondrashkina E, Kayatekin C, Matthews CR, Bilsel O. Microsecond acquisition of heterogeneous structure in the folding of a TIM barrel protein. *Proc Natl Acad Sci U S A*. 2008; 105:13367–13372. [PubMed: 18757725]
43. Kathuria SV, Guo L, Graceffa R, Barrea R, Nobrega RP, Matthews CR, Irving TC, Bilsel O. Minireview: structural insights into early folding events using continuous-flow time-resolved small-angle X-ray scattering. *Biopolymers*. 2011; 95:550–558. [PubMed: 21442608]
44. Hagen SJ, Eaton WA. Two-state expansion and collapse of a polypeptide. *J Mol Biol*. 2000; 301:1019–1027. [PubMed: 10966803]
45. Sinha KK, Udgaonkar JB. Dissecting the non-specific and specific components of the initial folding reaction of barstar by multi-site FRET measurements. *J Mol Biol*. 2007; 370:385–405. [PubMed: 17512542]
46. Sinha KK, Udgaonkar JB. 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; 353:704–718. [PubMed: 16188274]
47. Dasgupta A, Udgaonkar JB. Evidence for initial non-specific polypeptide chain collapse during the refolding of the SH3 domain of PI3 kinase. *J Mol Biol*. 2010; 403:430–445. [PubMed: 20837026]
The formation of a collapsed state within 150 μ s is demonstrated by FRET and additional spectroscopic methods. ANS binding to the collapsed state indicates early formation of hydrophobic clusters.
48. 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; 352:683–699. [PubMed: 16098987]
49. Ratner V, Sinev M, Haas E. Determination of intramolecular distance distribution during protein folding on the millisecond timescale. *J Mol Biol*. 2000; 299:1363–1371. [PubMed: 10873459]
50. Kimura T, Lee JC, Gray HB, Winkler JR. Site-specific collapse dynamics guide the formation of the cytochrome c' four-helix bundle. *Proc Natl Acad Sci U S A*. 2007; 104:117–122. [PubMed: 17179212]
51. Hamadani KM, Weiss S. Nonequilibrium Single Molecule Protein Folding in a Coaxial Mixer. *Biophys. J*. 2008; 95:352–365. [PubMed: 18339751]
52. Waldauer SA, Bakajin O, Ball T, Chen Y, Decamp SJ, Kopka M, Jager M, Singh VR, Wedemeyer WJ, Weiss S, et al. Ruggedness in the folding landscape of protein L. *HFSP J*. 2008; 2:388–395. [PubMed: 19436489]
53. Arai M, Iwakura M, Matthews CR, Bilsel O. Microsecond subdomain folding in dihydrofolate reductase. *J Mol Biol*. 2011; 410:329–342. [PubMed: 21554889]
54. Auton M, Holthauzen LM, Bolen DW. Anatomy of energetic changes accompanying urea-induced protein denaturation. *Proc Natl Acad Sci U S A*. 2007; 104:15317–15322. [PubMed: 17878304]
55. Bolen DW, Rose GD. Structure and energetics of the hydrogen-bonded backbone in protein folding. *Annu Rev Biochem*. 2008; 77:339–362. [PubMed: 18518824]
56. Tran HT, Mao A, Pappu RV. Role of backbone-solvent interactions in determining conformational equilibria of intrinsically disordered proteins. *J Am Chem Soc*. 2008; 130:7380–7392. [PubMed: 18481860]
57. 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; 103:12394–12399. [PubMed: 16894178]
58. Teufel DP, Johnson CM, Lum JK, Neuweiler H. Backbone-driven collapse in unfolded protein chains. *J Mol Biol*. 2011; 409:250–262. [PubMed: 21497607] FCS is used to probe the collapse of

unfolded proteins. The role of backbone interactions is highlighted by showing that a poly-glycine peptide, which is by definition devoid of side-chains, is nevertheless able to collapse in water.

59. England JL, Haran G. Role of solvation effects in protein denaturation: from thermodynamics to single molecules and back. *Annu Rev Phys Chem.* 2011; 62:257–277. [PubMed: 21219136] A review that describes the current status of our understanding of chemical denaturation, drawing on a large body of theoretical and experimental studies.
60. Ziv G, Haran G. Protein Folding, Protein Collapse, and Tanford's Transfer Model: Lessons from Single-Molecule FRET. *J Am Chem Soc.* 2009; 131:2942–2947. [PubMed: 19239269] A theoretical analysis of a series of single-molecule FRET studies of collapse, based on the coil-globule transition theory. A surprising connection is found between the free energy of collapse and the free energy of folding.
61. Grosberg AY, Kuznetsov DV. Quantitative Theory of the Globule-to-Coil Transition .1. Link Density Distribution in a Globule and Its Radius of Gyration. *Macromolecules.* 1992; 25:1970–1979.
62. Sanchez IC. Phase transition behavior of the isolated polymer chain. *Macromolecules.* 1979; 12:980–988.
63. Greene RF Jr, Pace CN. Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin. *J Biol Chem.* 1974; 249:5388–5393. [PubMed: 4416801]
64. Plaxco KW, Millett IS, Segel DJ, Doniach S, Baker D. Chain collapse can occur concomitantly with the rate-limiting step in protein folding. *Nat Struct Biol.* 1999; 6:554–556. [PubMed: 10360359]
65. Millett IS, Townsley LE, Chiti F, Doniach S, Plaxco KW. Equilibrium collapse and the kinetic 'foldability' of proteins. *Biochemistry.* 2002; 41:321–325. [PubMed: 11772031]
66. Jacob J, Dothager RS, Thiyagarajan P, Sosnick TR. Fully reduced ribonuclease A does not expand at high denaturant concentration or temperature. *J Mol Biol.* 2007; 367:609–615. [PubMed: 17292402]
67. Jacob J, Krantz B, Dothager RS, Thiyagarajan P, Sosnick TR. Early collapse is not an obligate step in protein folding. *J Mol Biol.* 2004; 338:369–382. [PubMed: 15066438]
68. Konuma T, Kimura T, Matsumoto S, Goto Y, Fujisawa T, Fersht AR, Takahashi S. Time-resolved small-angle X-ray scattering study of the folding dynamics of barnase. *J Mol Biol.* 2011; 405:1284–1294. [PubMed: 21146541]
69. Choy WY, Mulder FA, Crowhurst KA, Muhandiram DR, Millett IS, Doniach S, Forman-Kay JD, Kay LE. Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *J Mol Biol.* 2002; 316:101–112. [PubMed: 11829506]
70. Ziv G, Thirumalai D, Haran G. Collapse transition in proteins. *Phys Chem Chem Phys.* 2009; 11:83–93. [PubMed: 19081910]
71. Brooks CL 3rd. Protein and peptide folding explored with molecular simulations. *Acc Chem Res.* 2002; 35:447–454. [PubMed: 12069630]
72. Ptitsyn OB, Uversky VN. The molten globule is a third thermodynamical state of protein molecules. *FEBS Lett.* 1994; 341:15–18. [PubMed: 8137915]
73. Miller TF 3rd, Vanden-Eijnden E, Chandler D. Solvent coarse-graining and the string method applied to the hydrophobic collapse of a hydrated chain. *Proc Natl Acad Sci U S A.* 2007; 104:14559–14564. [PubMed: 17726097]
74. Sosnick TR, Barrick D. The folding of single domain proteins--have we reached a consensus? *Curr Opin Struct Biol.* 2011; 21:12–24. [PubMed: 21144739] A series of open questions in the protein folding field is discussed in this review, including the collapse transition, the diversity of folding pathways and the role of internal chain friction.

Figure 1. The collapse transition and protein folding

Unfolded proteins may change their configuration depending on the quality of the solution. In a good solvent they are expanded (top left) while in a bad solvent they are collapsed (top right). The collapse transition has been observed in both equilibrium and time-resolved experiments. The configuration of the chain may have an effect on folding thermodynamics. Thus, an expanded chain (whose radius of gyration, R_g , is much larger than that of the folded protein) is stabilized with respect to the folded state due to its higher configurational entropy. This is depicted by the schematic free energy surface on the left (U- unfolded state, F- folded state), plotted as a function of R_g . When the protein collapses, it loses much of its configurational entropy, and it is therefore less stable than the folded state- see the free energy surface on the right. The entropic loss accompanying collapse may thus facilitate folding.

Figure 2. Time-resolved SAXS experiments show collapse

R_g values for the protein dihydrofolate reductase, obtained from SAXS measurements following a fast jump in solution conditions [11]. The native and fully unfolded R_g values are also shown (in green and orange, respectively), to indicate that the SAXS experiment 'sees' a collapsed state whose size is intermediate between the two. This figure is reproduced with permission.

Figure 3. Sample calculation of the collapse free energy from single-molecule FRET measurements

The collapse free energy, $\Delta G_{U \rightarrow C}$ is shown in green triangles, and is found to be linear with denaturant concentration, with a slope that matches very well the folding transition m -value (black line). $\Delta G_{U \rightarrow C}$ is also decomposed in the figure into an enthalpic part (blue squares), and an entropic part (red circles). Note the substantial contribution of the configurational entropy (based on data and analysis in ref. [60]).