Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Where the complex things are: single molecule and ensemble spectroscopic investigations of protein folding dynamics

Satoshi Takahashi, Kiyoto Kamagata and Hiroyuki Oikawa

Progress in our understanding of the simple folding dynamics of small proteins and the complex dynamics of large proteins is reviewed. Recent characterizations of the folding transition path of small proteins revealed a simple dynamics explainable by the native centric model. In contrast, the accumulated data showed the substates containing residual structures in the unfolded state and partially populated intermediates, causing complexity in the early folding dynamics of small proteins. The size of the unfolded proteins in the absence of denaturants is likely expanded but still controversial. The steady progress in the observation of folding of large proteins has clarified the rapid formation of long-range contacts that seem inconsistent with the native centric model, suggesting that the folding strategy of large proteins is distinct from that of small proteins.

Address
Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, Miyagi 980-8577, Japan

Corresponding author: Takahashi, Satoshi (st@tagen.tohoku.ac.jp)

Current Opinion in Structural Biology 2016, **36**:1–9

This review comes from a themed issue on **Folding and binding**

Edited by **Ruth Nussinov** and **Jayant B Udgaonkar**

For a complete overview see the Issue and the Editorial

Available online 11th December 2015

http://dx.doi.org/10.1016/j.sbi.2015.11.006

Introduction

In the 1970s, Go and his collaborators enthusiastically investigated behaviors of a coarse-grained model of proteins using numerical calculations. When they placed a chain of 49 beads in a two-dimensional lattice and assumed attractive interactions only between pairs of beads forming native contacts, they observed cooperative transitions resembling heat denaturation curves of actual proteins (Figure 1a–c) [1]. The native centric model of proteins thereby originated with Go, but it was met with a lack of comprehension of many researchers in the 70s and 80s, leading him to propose the consistency principle of protein folding [2]. Now, the model is regarded as embodying the ''perfect funnel'' proposed in the energy landscape theory, which subsumes that the primary sequences of natural proteins are optimized for smooth folding transitions [3,4]. The native centric model incorporated into more sophisticated treatment of proteins reproduced not only the cooperative equilibrium transitions but also the tendency of folding rates and the structure of the folding transition state of small single domain proteins [5,6]. The energy landscape theory and the native centric model of proteins are milestones of current structural biology.

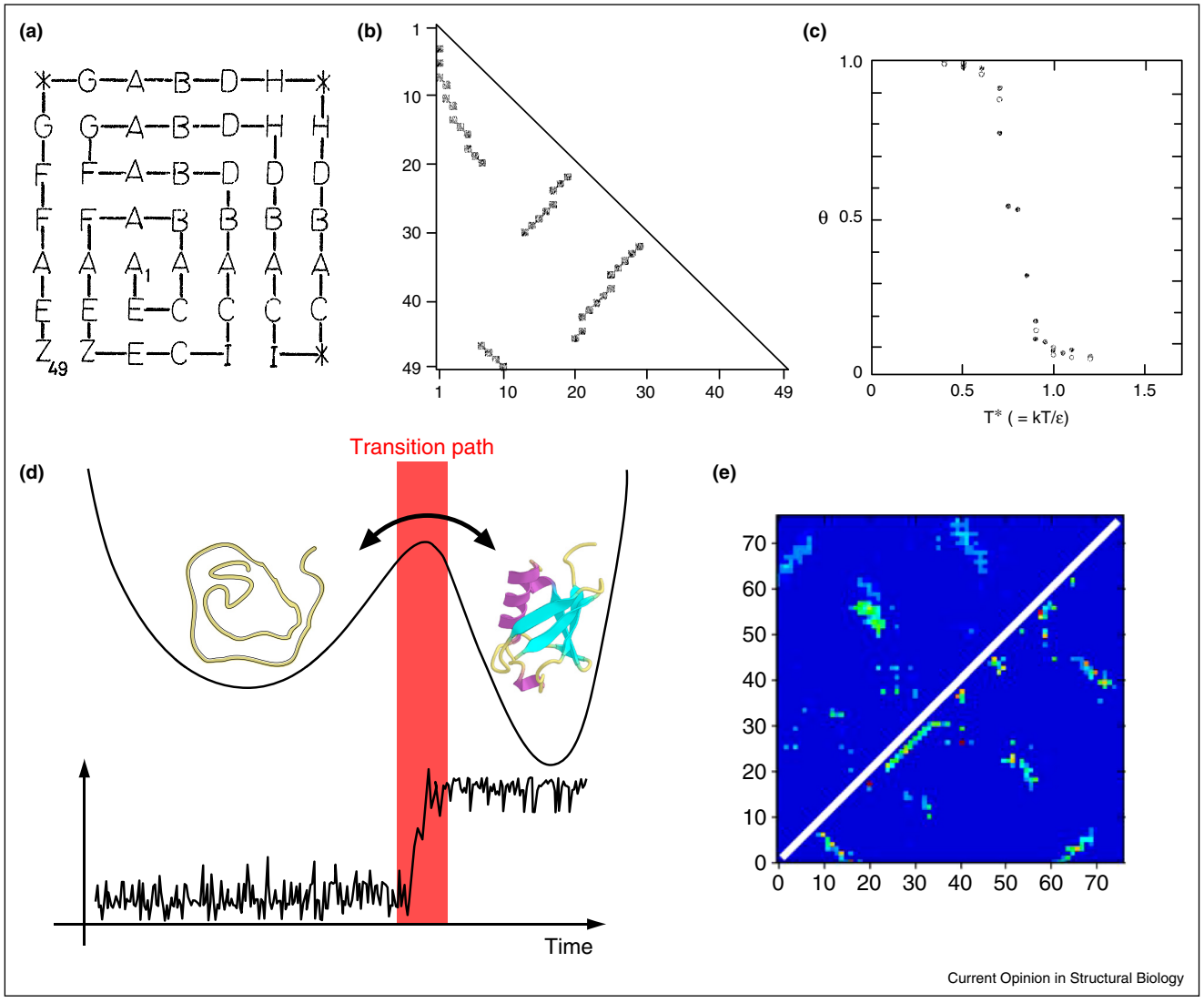
It is noteworthy, however, that important questions related to protein folding have remained unanswered [7,8]. The energy landscape theory and the native centric model are wonderfully simple, but neither includes consideration of details of intramolecular interactions arising from the primary sequence of amino acid residues and their interactions with water. Therefore, they do not explain how the primary sequences of natural proteins destabilize the misfolded conformations and minimize the roughness of the landscape. Partly because of the lack of such explanations, the *de-novo* design of artificial proteins and the prediction of large protein structures persist as difficult tasks. It is necessary to elucidate the mechanism that achieves apparent simplicity in the folding transition of small proteins and to reveal the complex folding of large proteins.

This review presents a summary of the exciting progress and notable discussions related to the simplicity and complexity of protein-folding dynamics that have appeared since 2013. We first introduce results demonstrating applicability of the native centric model for the dynamics of proteins in the transition path. We next introduce results showing structural and dynamical complexity in the unfolded state of even small proteins. Finally, we introduce progress in the investigations of complex folding mechanisms of large proteins.

Native centric model can explain the dynamics of proteins in the transition path

Small proteins with fewer than roughly 100 residues usually demonstrate two-state folding. They provided ample data supporting the energy landscape theory. The recent examination of the folding transition path of small proteins further demonstrated that the theory is almost perfectly satisfied. Chung and Eaton estimated the duration necessary for the folding transitions that are observed as discrete jumps in the single molecule fluorescence data, and termed the duration as the transition

Figure 1



Native centric potential can reproduce various properties of protein folding transitions. **(a)** Two-dimensional protein model originally proposed by Go *et al*. [1]. Each character represents one residue. **(b)** Contact map corresponding to the folded structure in the panel a. **(c)** Temperature denaturation transition calculated for the model protein in panel A assuming the native centric potential. The figures in panels (a)–(c) were adapted with permission from [1]. **(d)** One-dimensional reaction coordinate of protein folding. The two-state folding transition observed for small proteins can be described by one-dimensional potential diagram. The unfolded protein in the left can convert to the native state in one step. Thus, in the single molecule time series measurements, the transition can be detected as the jumps in the fluorescence signals as shown in the putative time series shown in the bottom. Time duration necessary for the folding transition is designated as the transition path time. In the native basin, the folded structure of ubiquitin is presented. **(e)** The contact map in the lower right half represents the native structure of ubiquitin. The upper left half represents the probability of forming inter-residue contacts in the transition path of the all atom MD calculation for folding of ubiquitin. The data were adapted from [15•].

path time (Figure 1d) [9,10••,11•]. The duration was in the range of one to several tens of microseconds, which is consistent with the speed limit of protein folding. All-atom molecular dynamics (MD) calculations of protein folding roughly reproduced the duration [12,13•,14]. The transition path time can be analyzed using Kramer's theory assuming a diffusive barrier crossing over a single reaction coordinate. Furthermore, the native centric potential can explain the behaviors of proteins in the transition path of the all-atom MD calculations (Figure 1e) [15•,16•]. Consequently, the folding transition occurs as a simple diffusive process over a one-dimensional reaction coordinate with a minimum contribution from the non-native contacts.

Another example suggesting a small contribution of the non-native contacts in the transition state of small proteins is identification of the origin of internal friction. The internal friction can be estimated by examining the folding rates at different solvent viscosities and reflects the roughness of the energy landscape along the transition paths [17]. The internal friction for the folding of α-helical proteins is greater than that for β or α/β proteins [17]. Best and his collaborators concluded that the friction arises from the changes in the dihedral angles of polypeptide backbone [18••,19]. The friction becomes greater for helical proteins because the unfolded proteins mainly possess dihedral angles near beta region. The conclusion suggests that the frustration caused by the transformations of secondary structures is more important than the possible frustration by intrachain non-native contacts for the transition state of small proteins.

In accordance with the simplicity of the transition path, it is usually difficult to find evidence for the inhomogeneity in the folding transitions of small proteins [20]. For example, homologous proteins L and G were thought to form N-terminal and C-terminal turns, respectively, in the transition state; however, a recent reinvestigation by Sosnick *et al.* revealed that both turns are similarly formed in the transition state for both proteins [21•]. Maqusee *et al.* reported that different unfolding routes are undertaken in the folding transition state of a small protein, SH3, with small changes in the unfolding conditions or mutations [22•]. Accordingly, while the important exception exists, the folding transition state of small proteins is highly homogeneous and is strongly controlled by the native structures.

Structural and dynamical properties of the unfolded state of small proteins

Sizes of unfolded proteins

Past investigations have established that the unfolded proteins in the presence of high concentrations of denaturants are expanded. Their radius of gyration shown against the number of residues obeys the scaling law that is expected for a self-avoiding random coil [23]. In contrast, the size of a more important state, the unfolded state in the absence of denaturant, remains controversial. Because the unfolded proteins in the all-atom MD calculations are too collapsed, the accurate size of the unfolded state is important for the evaluation of force fields [14,24–26].

Single-molecule fluorescence resonance energy transfer (smFRET) studies have suggested that the unfolded state of proteins such as protein L in the absence of denaturant become smaller than those in the presence of denaturant [27]. Compaction occurs gradually as the denaturant concentration decreases, suggesting the absence of an energy barrier between the expanded and collapsed conformations [25]. Schuler *et al.* observed that unfolded proteins at the lower concentration of denaturant collapse moderately and approximate the scaling law expected for Gaussian polymers in Θ solvents [28•]. In contrast, small-angle X-ray scattering (SAXS) has demonstrated that the radius of gyration of unfolded proteins is expanded immediately after dilution of the denaturant [29,30]. The detection of the expanded state in kinetic experiments means that there exists an energy barrier separating the expanded and the collapsed states.

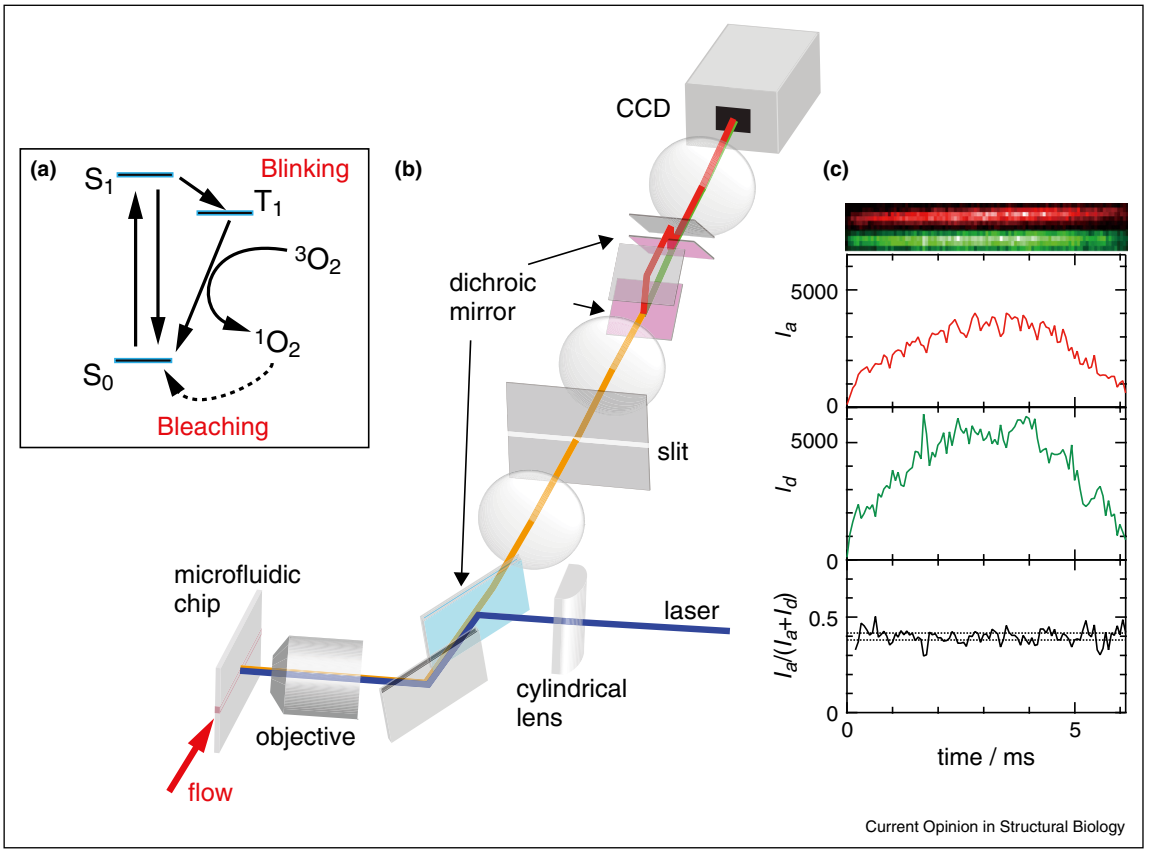
Several recent results rather supported the conclusion from the scattering measurements. Plaxco *et al.* examined the size of polyethylene glycol (PEG) using smFRET and small angle neutron scattering [31•]. Although the smFRET data of the doubly labeled PEG showed an increase of the FRET efficiency occurring concomitantly with the decrease in the denaturant concentration, the neutron scattering data showed that the unlabeled PEG remains expanded irrespective of the denaturant concentration. Although the origin of the discrepancy was not explained, it is clear that smFRET tends to estimate smaller sizes at the lower concentration of denaturant. Cytochrome c_{552} possesses an expanded unfolded state after dilution of the denaturant, which is consistent with the presence of the energy barrier for the collapse [32]. Similar barrier-limited collapse with a time constant of approximately 30 μs was also observed for cytochrome *c* [33••].

The possibility still remains that unfolded proteins very close to zero denaturant concentration contract to some degree, as demonstrated in the reduced scaling exponent by sm-FRET [28•]. In fact, a small collapse of barnase was observed using time-resolved SAXS [34]. To resolve the controversy, however, this issue should be investigated by methods other than sm-FRET or SAXS. The limited time resolution of smFRET (approx. 1 ms) as well as that of SAXS (limited by the mixing device) might average the subpopulations in the denatured ensemble [33••]. We explain below that the unfolded proteins are not a perfect random coil and that they possess substates with residual structures. Further analysis based on methods that can resolve the substates is required.

Heterogeneity and residual structures of the unfolded proteins

Munoz *et al.* first improved the time resolution of the smFRET measurements from ca. 1 ms to less than 100 μs [35]. We achieved a similar time resolution based on a different approach: line confocal microscopy (Figure 2) [36•]. We demonstrated that the unfolded state of the B domain of protein A (BdpA), a small three-helix bundle assumed to be a two-state protein, comprises several substates (Figure 3) [37••]. A gradual expansion of the native state as the increase of the denaturant was also detected. Accordingly, the increased temporal and structural resolution of the smFRET measurements

Figure 2

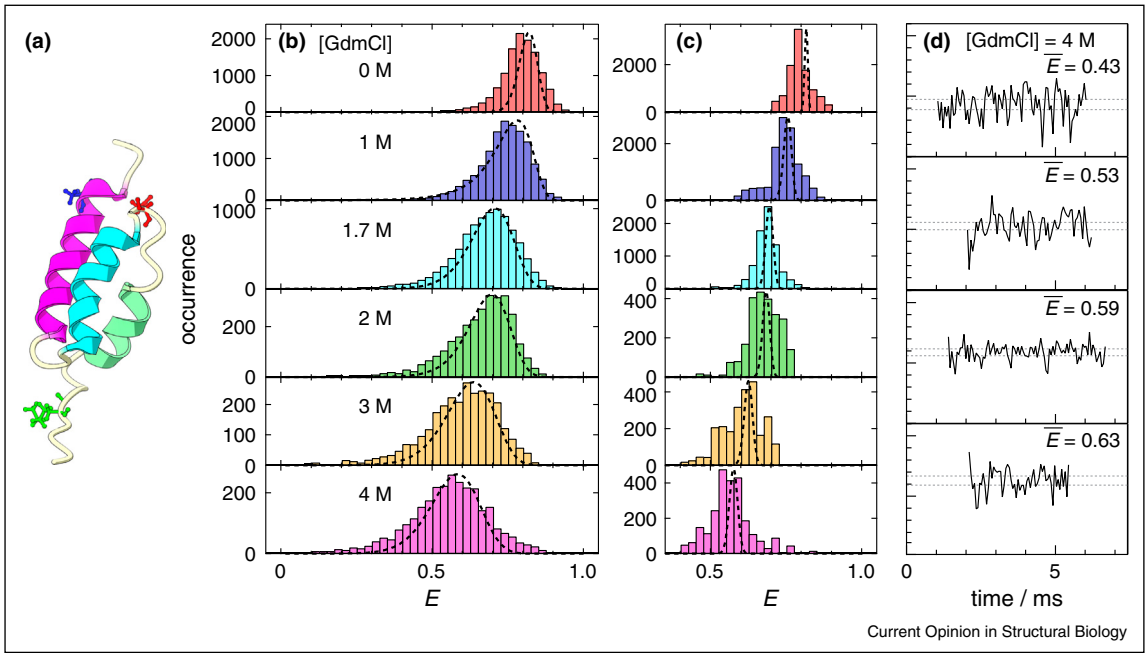


(a) Photocycle of organic fluorophores. For single-molecule fluorescence measurements, the fast photocycle between S_0 and S_1 is essential. However, the photocycle is interrupted occasionally by the formation of T_1, causing blinking and bleaching of the fluorescence signals. The dissolved oxygen, $^3O_2$, quenches T_1 to S_0, and can restarts the photocycle; however, the $^3O_2$ concentration is usually minimized to prevent blinking and bleaching in the conventional confocal microscopy. In contrast, the line confocal microscopy flows a sample solution with dissolved $^3O_2$ rapidly. The strategy enables an increase of the photon count rate for more than several orders of magnitude larger than the conventional method. **(b)** Schematic diagram of the line confocal microscopy for single-molecule FRET efficiency measurements. The doubly labeled sample was introduced to the flow cell at the fast flow speed. The excitation laser is focused linearly along the flow channel, so that flowing molecules can be excited continuously. The fluorescence photons from the donor and acceptor fluorophors are collected by the objective, separated spatially by dichroic mirrors, and imaged using CCD. **(c)** Example of the data obtained by the system. Top image shows artificially colored image of a flowing molecule detected by CCD. The second and third images represent the fluorescence intensities for the acceptor (I_a) and donor (I_d) estimated from the top image. The bottom trace represents the apparent FRET efficiency time series estimated by dividing I_a with $I_a + I_d$.

revealed the deviation of BdpA from the two-state folding.

Increasingly, accumulated data have shown deviation from the two-state folding for small proteins that were known previously as the two-state proteins. Based on temperature-jump infrared spectroscopy, a partial formation of helices in the nanosecond time domain and the subsequent growth of helices in the microsecond time domain were observed for BdpA [38]. A similar observation was reported for leucine zipper dimer based on pH-jump infrared spectroscopy [39•]. Careful examination of the equilibrium unfolding of gpW by nuclear magnetic resonance spectroscopy and all atom MD calculation demonstrated that some residues unfold in a two-state manner and others in a three-state manner [40]. For SH3, the deviation from the two-state folding and the formation of two intermediates including dry molten globule were discussed [41,42]. The initial folding phase of Acl-CoA binding protein (ACBP) was inferred as occurring heterogeneously based on the comparison of the all atom MD calculations and the rapid mixing fluorescence results [43]. These results demonstrate that the apparent two-state behavior of small proteins is achieved as a result of the complex dynamics involving substates in the unfolded state and intermediates accumulated in a small fraction, which often possess a fragment of secondary structures promoting the formation of non-local contacts.

Figure 3



Equilibrium unfolding transition of the B domain of protein A (BdpA) observed by the line confocal miscroscopy. **(a)** Native structure of BdpA. Residues 5 colored green and 55 colored blue were labeled by fluorophores. **(b)** FRET efficiency histograms of BdpA of the raw data obtained at 180 μs time resolution. Dotted lines represent the noise width estimated from the average photon numbers of signals and backgrounds assuming single FRET efficiency. **(c)** The same FRET efficiency histograms after the moving average of the time series data of 1 ms. Although the data in panel B show a gradual shift of the unfolded state efficiency, the data after time averaging show a gradual shift of the native peak observed at 0–2 M Gdm, and splitting of the unfolded state distributions. Dotted lines represent the noise width after the time averaging. The observed distributions for the unfolded state are broader than the noise width, showing the heterogeneity of the unfolded state. **(d)** Examples of the FRET efficiency time series obtained at 4-M GdmCl. Individual traces showed different efficiencies during the observation time of a few milliseconds. All the figures were adapted with permission from [37••]. Copyright 2015 American Chemical Society.

The presence of substates in the unfolded state even for small single-domain proteins is related to the residual structures. The double mutant cycle and the paramagnetic relaxation measurements for the unfolded state of L9 demonstrated residual structures with non-native long-range contacts [44••]. Residual structures in the unfolded state of ACBP have been reported [45]. The presence of partially collapsed hydrophobic clusters having a microsecond lifetime was also suggested in all-atom MD calculations [13•]. Designed proteins with very similar sequences but having different folded structures possess distinct structures in the very early stage of folding likely in the unfolded state [46]. These results support the view that the unfolded state of small proteins is not a simple random coil but is instead a mixture of heterogeneous residual structures.

As a final remark, we describe very slow fluctuations suggested for BdpA [37••]. The individual smFRET traces for the unfolded BdpA showed constant but different efficiencies for several milliseconds (Figure 3c). That observation suggests slow fluctuations and contrasts against the fast fluctuations reported for many unfolded proteins. Fluorescence correlation measurements for cold shock protein revealed reconfiguration time of ~100 ns [47]. The time scale for the contact formation was 0.1 μs for a 6 residue loop and more than 20 μs for a 66 residue loop [48]. Our results might be consistent with putative slow fluctuations inferred for the unfolded state of protein L [49]. Further investigations must be conducted to clarify the origin of the slow fluctuations.

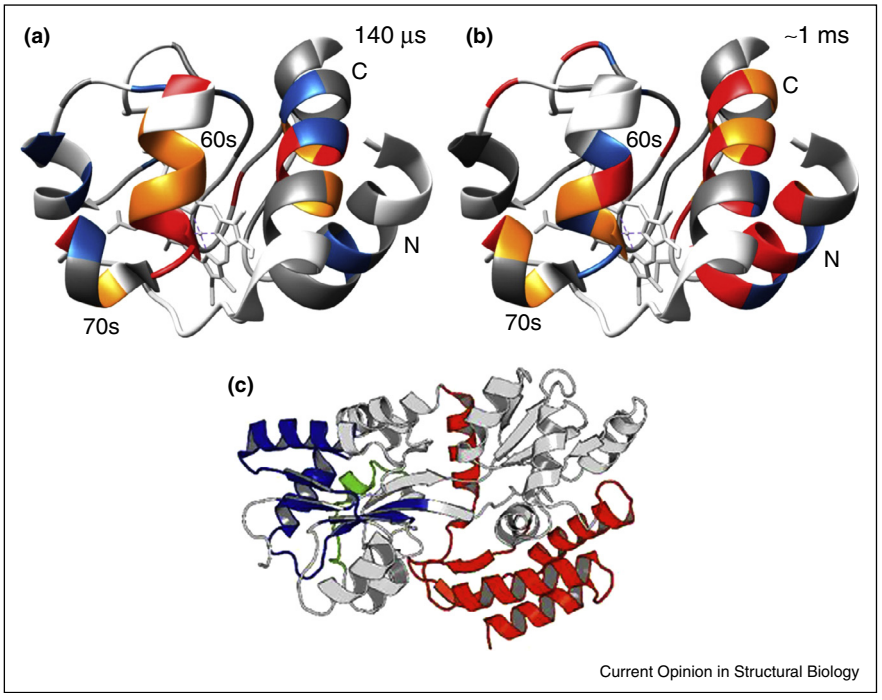
Complex folding dynamics of proteins larger than 100 residues

Single domain proteins larger than roughly 100 residues generally form folding intermediates and demonstrate folding dynamics distinct from small proteins [50]. We observed previously that unfolded proteins such as cytochrome *c* and apomyoglobin first collapse to the intermediate within submilliseconds, followed by slower conversion to a native state [51,52]. Many native and some non-native structures are formed in the collapse. In particular, the long-range contacts between the N- and C-termini are likely formed in the intermediates, implying that the collapse entails the formation of an important framework of the native structure [53].

Recent results began to resolve events occurring in the collapse for large single-domain proteins. Roder *et al.* developed a device for pulsed hydrogen deuterium exchange with time resolution of 100 μs [54••]. They observed that the fragments of C-terminal and 60s helices of cytochrome *c*, which are close in sequence, associate before the contact formation between the N- and C-terminal helices (Figure 4a and b). The initial formation of the local structures leading to the long-range contacts strongly suggests that the fast collapse is designed in the arrangement of the hydrophobic residues. Bilsel *et al.* observed that the collapse of cytochrome *c* occurs as a barrier-crossing event [33••]. Based on the time correlation of single-molecule fluorescence without reliance on the kinetic triggering of folding, several intermediates that connect the unfolded and native states linearly were resolved in the folding of yeast cytochrome *c* [55••]. The lifetime of the intermediate of a mutant of apomyoglobin, in which one non-native interaction was eliminated, was not significantly altered from that of the wild type, suggesting the presence of further non-native contacts in the intermediate [56]. Finally, the non-native helix formation in the folding of β-lactoglobulin occurs concomitantly with the rapid collapse, and was proposed to promote the long-range contact formation [57]. These results suggest that the initial collapse of large single-domain proteins is designed to occur rapidly and to form the long-range native contacts together with the non-native contacts.

The rapid formation of the long-range contacts might not be consistent with the native centric model. This becomes more apparent for the folding of some large multidomain proteins. Englander *et al.* developed a mass spectroscopic method to detect the exchange of amide protons of large proteins at the residue level, and clarified that the different domains of maltose binding protein form with different timing (Figure 4c) [58••]. The fast folding domain possesses the long-range contacts separated more than 250 residues. The folding of different loops of adenylate kinase occurs with different timing and the loop encompassing residues 1 to 88 forms much faster than the small loop between residues 121 and 155 [59•]. The folding of repeat proteins occurs frequently from one domain having the greatest stability [60•]. These results demonstrate that different domains of multidomain

Figure 4



Stepwise collapse observed for cytochrome *c*. Results of the ultrafast pulsed hydrogen deuterium exchange experiments for the folding process of cytochrome *c* at 140 μs **(a)** and 1 ms **(b)** after initiation of the folding reaction. Regions colored red correspond to the amide protons protected. At 140 μs, the small clusters involving the C-terminal and 60s helices are formed. However, the N-terminal helix is not protected. In contrast, both of the N- and C-terminal helices are strongly protected at 1 ms, suggesting the formation of the long range contact between the two helices. The figures were adapted with permission from [54••]. Copyright 2013 American Chemical Society. **(c)** The stepwise folding of a large multidomain protein, maltose binding protein, examined by the hydrogen deuterium exchange pattern of kinetic mass spectroscopy. The blue region includes residues 9–43, 60–62 and 260–278 and is protected with time constant of ~7 s. The red region includes residues 180–195 and 315–370 and is protected with time constant of ~100 s. The green region includes residues 78–89 and is protected slightly faster than the red region.The figure was adapted from [58••].

proteins fold with different timing, which likely deviates from the behaviors suggested by the native centric model. In contrast, recent analyses of the folding of dihydrofolate reductase, which comprises two domains and shows complex folding kinetics involving six phases, revealed that the native centric model can reproduce the folding kinetics for the wild type and the circular permutated mutant [61,62•]. Accordingly, in addition to the native centric behavior, the large proteins seem to have developed the distinct strategy to fold their complex structures.

Conclusion

We first discussed the applicability of the native centric model for the dynamics of small proteins in the folding transition path, and the contrasting complexity in the unfolded state and the deviation from the two-state folding. The simplicity in the folding transition and the contrasting complexity in the unfolded state strongly suggest the existence of a mechanism that minimizes the contribution of the non-native contact in the folding transition state of small proteins. We next discussed that the folding of large proteins involves the rapid formation of the intermediates possessing localized domains with long-range contacts, suggesting the existence of a mechanism deviating from the native centric model which is required for the formation of large structures. We propose that the native centric behavior is one strategy chosen by nature to fold small proteins, and that additional strategies were developed to fold larger proteins. Thus, the complex things are mainly in the unfolded state for small proteins and in the entire folding process for large proteins.

Go proposed that local structures with native and non-native structures might form in the early folding stages; however, only the native structures can grow and overcome the folding transition state, because native local structure is structurally consistent with other native local structures [2]. The accumulated data demonstrating the residual structures might be consistent with the proposal; however, the information how the substates and residual structures are connected each other and how they grow to form the transition state is mostly unknown. Further characterizations of the unfolded state as well as the folding transition path based on the advanced methods of single molecule and ensemble spectroscopies are urgently necessary. The theoretical analysis of single-molecule data might be promising to clarify the dynamics [63,64]. The advanced spectroscopic methods should further be designed and applied to characterize the folding mechanism of each of large proteins to understand their design strategies.

Acknowledgements

This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan to S.T. and K.K.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Takemori H, Ueda Y, Go N: **Studies on protein folding, unfolding and fluctuations by computer simulation. The effect of specific amino acid sequence represented by specific inter-unit interactions**. *Int J Pept Prot Res* 1975, **7**:445-459.
2. Go N: **Theoretical studies of protein folding**. *Ann Rev Biophys Bioeng* 1983, **12**:183-210.
3. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG: 1: **Funnels, pathways and energy landscape of protein folding: a synthesis**. *Proteins* 1995, **21**:167-195.
4. Onuchic JN, Wolynes PG: **Theory of protein folding**. *Curr Opin Struct Biol* 2004, **14**:70-75.
5. Muñoz V: **A simple theoretical model goes a long way in explaining complex behavior in protein folding**. *Proc Nat Acad Sci USA* 2014, **111**:15863-15864.
6. Hills RD Jr, Brooks CL 3rd: **Insights from coarse-grained Go models for protein folding and dynamics**. *Int J Mol Sci* 2009, **10**:889-905.
7. Dill KA, MacCallum JL: **The protein-folding problem, 50 years on**. *Science* 2012, **338**:1042-1046.
8. Englander SW, Mayne L: **The nature of protein folding pathways**. *Proc Natl Acad Sci USA* 2014, **111**:15873-15880.
9. Chung HS, Piana-Agostinetti S, Shaw DE, Eaton WA: **Structural origin of slow diffusion in protein folding**. *Science* 2015, **349**:1504-1510.
10. •• Chung HS, Eaton WA: **Single-molecule fluorescence probes dynamics of barrier crossing**. *Nature* 2013, **502**:685-688.
11. • Chung HS, McHale K, Louis JM, Eaton WA: **Single-molecule fluorescence experiments determine protein folding transition path times**. *Science* 2012, **335(6071)**:981-984.
The article demonstrated that the transition path time of protein folding can be quantitated, and is the starting point for the exciting discussions on the key dynamical event.
12. Piana S, Lindorff-Larsen K, Shaw DE: **Protein folding kinetics and thermodynamics from atomistic simulation**. *Proc Natl Acad Sci USA* 2012, **109**:17845-17850.
13. • Piana S, Lindorff-Larsen K, Shaw DE: **Atomic-level description of ubiquitin folding**. *Proc Natl Acad Sci USA* 2013, **110**:5915-5920.
14. Piana S, Klepeis JL, Shaw DE: **Assessing the accuracy of physical models used in protein-folding simulations: quantitative evidence from long molecular dynamics simulations**. *Curr Opin Struct Biol* 2014, **24**:98-105.
15. • Best RB, Hummer G, Eaton WA: **Native contacts determine protein folding mechanisms in atomistic simulations**. *Proc Natl Acad Sci U S A* 2013, **110**:17874-17879.
During the transition path in all atom MD results of natural proteins, the probability of forming native contacts is larger than that of the non-native contact. However, an artificial protein, α_3D, possesses significant contributions of the non-native contacts.
16. • Henry ER, Best RB, Eaton WA: **Comparing a simple theoretical model for protein folding with all-atom molecular dynamics simulations**. *Proc Natl Acad Sci USA* 2013, **110**:17880-17885.
17. Hagen SJ: **Solvent viscosity and friction in protein folding dynamics**. *Curr Prot Pept Sci* 2010, **11**:385-395.
18. •• de Sancho D, Anshul S, Best RB: **Molecular origins of internal friction effects on protein-folding rates**. *Nat Commun* 2014, **5**:4307.
19. Zheng W, De Sancho D, Hoppe T, Best RB: **Dependence of internal friction on folding mechanism**. *J Am Chem Soc* 2015, **137**:3283-3290.

20. Shandiz AT, Baxa MC, Sosnick TR: **A "Link-Psi" strategy using crosslinking indicates that the folding transition state of ubiquitin is not very malleable**. *Prot Sci* 2012, **21**:819-827.

21. • Baxa MC, Yu W, Adhikari AN, Ge L, Xia Z, Zhou R, Freed KF, Sosnick TR: **Even with nonnative interactions, the updated folding transition states of the homologs Proteins G & L are extensive and similar**. *Proc Natl Acad Sci USA* 2015, **112**:8302-8307.

22. • Guinn EJ, Jagannathan B, Marqusee S: **Single-molecule chemo-mechanical unfolding reveals multiple transition state barriers in a small single-domain protein**. *Nat Commun* 2015, **6**:6861.

23. Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, Dothager RS, Seifert S, Thiyagarajan P, Sosnick TR, Hasan MZ, Pande VS, Ruczinski I, Doniach S, Plaxco KW: **Random-coil behavior and the dimensions of chemically unfolded proteins**. *Proc Natl Acad Sci USA* 2004, **101**:12491-12496 Erratum in: *Proc Natl Acad Sci USA* 2005, 102:14475..

24. Skinner JJ, Yu W, Gichana EK, Baxa MC, Hinshaw JR, Freed KF, Sosnick TR: **Benchmarking all-atom simulations using hydrogen exchange**. *Proc Natl Acad Sci USA* 2014, **111**:15975-15980.

25. Best RB, Zheng W, Mittal J: **Balanced protein-water interactions improve properties of disordered proteins and non-specific protein association**. *J Chem Theory Comput* 2014, **10**:5113-5124.

26. Piana S, Donchev AG, Robustelli P, Shaw DE: **Water dispersion interactions strongly influence simulated structural properties of disordered protein States**. *J Phys Chem B* 2015, **119**:5113-5123.

27. Haran G: **How, when and why proteins collapse: the relation to folding**. *Curr Opin Struct Biol* 2012, **22(1)**:14-20.

28. • Hofmann H, Soranno A, Borgia A, Gast K, Nettels D, Schuler B: **Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy**. *Proc Natl Acad Sci USA* 2012, **109**:16155-16160.

29. 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.

30. Yoo TY, Meisburger SP, Hinshaw J, Pollack L, Haran G, Sosnick TR, Plaxco K: **Small-angle X-ray scattering and single-molecule FRET spectroscopy produce highly divergent views of the low-denaturant unfolded state**. *J Mol Biol* 2012, **418**:226-236.

31. • Watkins HM, Simon AJ, Sosnick TR, Lipman EA, Hjelm RP, Plaxco KW: **Random coil negative control reproduces the discrepancy between scattering and FRET measurements of denatured protein dimensions**. *Proc Natl Acad Sci USA* 2015, **112**:6631-6636.

32. Yamada S, Bouley Ford ND, Keller GE, Ford WC, Gray HB, Winkler JR: **Snapshots of a protein folding intermediate**. *Proc Natl Acad Sci USA* 2013, **110**:1606-1610.

33. •• Kathuria SV, Kayatekin C, Barrea R, Kondrashkina E, Graceffa R, Guo L, Nobrega RP, Chakravarthy S, Matthews CR, Irving TC, Bilsel O: **Microsecond barrier-limited chain collapse observed by time-resolved FRET and SAXS**. *J Mol Biol* 2014, **426**:1980-1994.

This article demonstrates that the kinetic CD, SAXS and fluorescence lifetime data based on the continuous-flow rapid mixing device having mixing time of several tens of microseconds can now be routinely obtained.

34. 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.

35. Campos LA, Liu J, Wang X, Ramanathan R, English DS, Muñoz V: **The photoprotection strategy for microsecond-resolution single-molecule fluorescence spectroscopy**. *Nat Methods* 2011, **8**:143-146.

36. • Oikawa H, Suzuki Y, Saito M, Kamagata K, Arai M, Takahashi S: **Microsecond dynamics of an unfolded protein by a line confocal tracking of single molecule fluorescence**. *Sci Rep* 2013, **3**:2151.

37. •• Oikawa H, Kamagata K, Arai M, Takahashi S: **Complexity of the folding transition of the B domain of protein A revealed by the high-speed tracking of single-molecule fluorescence time series**. *J Phys Chem B* 2015, **119**:6081-6091.

38. Davis CM, Cooper AK, Dyer RB: **Fast helix formation in the B domain of protein A revealed by site specific infrared probes**. *Biochemistry* 2015, **54**:1758-1766.

39. • Donten ML, Hassan S, Popp A, Halter J, Hauser K, Hamm P: **pH-Jump induced leucine zipper folding beyond the diffusion limit**. *J Phys Chem B* 2015, **119**:1425-1432.

In contrast to the temperature jump method that can monitor only the unfolding processes, the pH jump method became matured and was applied to investigate the forward folding processes.

40. Sborgi L, Verma A, Piana S, Lindorff-Larsen K, Cerminara M, Santiveri CM, Shaw DE, de Alba E, Munoz V: **Interaction networks in protein folding via atomic-resolution experiments and long-time-scale molecular dynamics simulations**. *J Am Chem Soc* 2015, **137**:6506-6516.

41. Kishore M, Krishnamoorthy G, Udgaonkar JB: **Critical evaluation of the two-state model describing the equilibrium unfolding of the PI3K SH3 domain by time-resolved fluorescence resonance energy transfer**. *Biochemistry* 2013, **52**:9482-9496.

42. Dasgupta A, Udgaonkar JB, Das P: **Multistage unfolding of an SH3 domain: an initial urea-filled dry molten globule precedes a wet molten globule with non-native structure**. *J Phys Chem B* 2014, **118**:6380-6392.

43. Voelz VA, Jäger M, Yao S, Chen Y, Zhu L, Waldauer SA, Bowman GR, Friedrichs M, Bakajin O, Lapidus LJ, Weiss S, Pande VS: **Slow unfolded-state structuring in Acyl-CoA binding protein folding revealed by simulation and experiment**. *J Am Chem Soc* 2012, **134**:12565-12577.

44. •• Cho JH, Meng W, Sato S, Kim EY, Schindelin H, Raleigh DP: **Energetically significant networks of coupled interactions within an unfolded protein**. *Proc Natl Acad Sci USA* 2014, **111**:12079-12084.

The contact energy of the long-range interactions present in the unfolded state was quantitated. The interaction might impact the interpretation of Φ-values.

45. Ozenne V, Noel JK, Heidarsson PO, Brander S, Poulsen FM, Jensen MR, Kragelund BB, Blackledge M, Danielsson J: **Exploring the minimally frustrated energy landscape of unfolded ACBP**. *J Mol Biol* 2014, **426**:722-734.

46. Giri R, Morrone A, Travaglini-Allocatelli C, Jemth P, Brunori M, Gianni S: **Folding pathways of proteins with increasing degree of sequence identities but different structure and function**. *Proc Natl Acad Sci USA* 2012, **109**:17772-17776.

47. Nettels D, Gopich IV, Hoffmann A, Schuler B: **Ultrafast dynamics of protein collapse from single-molecule photon statistics**. *Proc Natl Acad Sci USA* 2007, **104**:2655-2660.

48. Bouley Ford ND, Shin DW, Gray HB, Winkler JR: **Intrachain contact dynamics in unfolded cytochrome cb_{562}**. *J Phys Chem B* 2013, **117**:13206-13211.

49. 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 USA* 2007, **104**:1528-1533.

50. Takahashi S, Kamagata K: **Staring at a protein: ensemble and single-molecule investigations on protein-folding dynamics**. *Adv Chem Phys* 2011, **146**:3-22.

51. 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.

52. 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 USA* 2004, **101**:1171-1176.

53. Uzawa T, Nishimura C, Akiyama S, Ishimori K, Takahashi S, Dyson HJ, Wright PE: **Hierarchical folding mechanism of apomyoglobin revealed by ultra-fast H/D exchange coupled with 2D NMR**. *Proc Natl Acad Sci USA* 2008, **105**:13859-13864.

54. •• Fazelinia H, Xu M, Cheng H, Roder H: **Ultrafast hydrogen exchange reveals specific structural events during the initial stages of folding of cytochrome *c***. *J Am Chem Soc* 2014, **136**:733-740.
A microfluidic device was developed to achieve the ultrarapid and multiple mixings required for the pulsed hydrogen deuterium exchange experiments, and clarified the initial collapse of cytochrome *c*.

55. •• Otosu T, Ishii K, Tahara T: **Microsecond protein dynamics observed at the single-molecule level**. *Nat Commun* 2015, **6**:7685.
A drastically new approach was applied to resolve the network of jumps among different intermediates of cytochrome *c* without using kinetic triggering of the folding reaction.

56. Aoto PC, Nishimura C, Dyson HJ, Wright PE: **Probing the non-native H helix translocation in apomyoglobin folding intermediates**. *Biochemistry* 2014, **53**:3767-3780.

57. Konuma T, Sakurai K, Yagi M, Goto Y, Fujisawa T, Takahashi S: **Highly collapsed conformation of the initial folding intermediates of β-lactoglobulin with Non-Native α-Helix**. *J Mol Biol* 2015, **427**:3158-3165.

58. •• Walters BT, Mayne L, Hinshaw JR, Sosnick TR, Englander SW: **Folding of a large protein at high structural resolution**. *Proc Natl Acad Sci USA* 2013, **110**:18898-18903.
Application of the mass-spectroscopic analysis of the hydrogen-deuterium exchange for a very large protein. The method can avoid the aggregate formation observed for large proteins by conducting the experiments at the protein concentration of less than 1 μM.

59. • Orevi T, Ben Ishay E, Gershanov SL, Dalak MB, Amir D, Haas E: **Fast closure of N-terminal long loops but slow formation of β strands precedes the folding transition state of *Escherichia coli* adenylate kinase**. *Biochemistry* 2014, **53**:3169-3178.

60. • Dao TP, Majumdar A, Barrick D: **Highly polarized C-terminal transition state of the leucine-rich repeat domain of PP32 is governed by local stability**. *Proc Natl Acad Sci USA* 2015, **112**:E2298-E2306.

61. Arai M, Iwakura M, Matthews CR, Bilsel O: **Microsecond subdomain folding in dihydrofolate reductase**. *J Mol Biol* 2011, **410**:329-342.

62. • Inanami T, Terada TP, Sasai M: **Folding pathway of a multidomain protein depends on its topology of domain connectivity**. *Proc Natl Acad Sci USA* 2014, **111**:15969-15974.

63. Taylor JN, Li CB, Cooper DR, Landes CF, Komatsuzaki T: **Error-based extraction of states and energy landscapes from experimental single-molecule time-series**. *Sci Rep* 2015, **5**:9174.

64. Matsunaga Y, Kidera A, Sugita Y: **Sequential data assimilation for single-molecule FRET photon-counting data**. *J Chem Phys* 2015, **142**:214115.