WATCHING SINGLE BIOLOGICAL MOLECULES FOLD USING LASER TWEEZERS

BY MICHAEL T. WOODSIDE

SUMMARY

Using optical tweezers to unravel structures in biological molecules, we can watch single molecules fold directly, providing new physical insight into this crucial biological process.

THE FOLDING PROBLEM AND ENERGY LANDSCAPES

One of the remarkable features of biological molecules like proteins, DNA, and RNA is that they can form intricate three-dimensional structures. These molecules are alike in being linear polymers, consisting of simple chains of amino acids (proteins) or nucleotides (DNA and RNA). After synthesis in the cell, however, these chains "fold" into a vast diversity of shapes, from the elegant simplicity of the DNA double helix to the magnificent complexity of the ribosome, the enzyme that makes all proteins. Such structural diversity is the essential feature that underlies the functional versatility of these molecules, because of the tight link between structure and function. Understanding how biomolecules fold is critical: each molecule must fold into the correct structure to ensure correct function, and "misfolding" (forming the wrong structure) is linked to a wide variety of diseases [1].

The main challenge in solving the folding problem is not our understanding of the interactions within the polymer chain and between the polymer chain and its aqueous medium that drive folding, but rather their combinatorial complexity. A classic thought-experiment from Cyrus Levinthal showed that even for a small protein, a random search for the correct structure would take trillions of universe lifetimes, in contrast to the typical folding timescale of milliseconds to minutes [2]. The resolution to "Levinthal's paradox" is that folding does not involve a random search—instead, the search is biased by the favorable energies of native (correct) interactions within the chain as they form during folding. The modern picture of folding, arising from the theory of glasses, views folding as a search over a 'landscape' representing the energy of the molecule as a function of all possible configurations, where the search is driven by thermal diffusion within the conformational space [3]. This landscape is typically funnel-shaped (Fig. 1A), guiding the search rapidly and reliably to the native structure. Understanding how structure forms thus reduces to understanding the shape of the energy landscape and the paths taken across it during folding.

A key feature of folding landscapes is that there are usually energy barriers that must be crossed before reaching the native state. Such barriers arise from the delicate balance between enthalpy and entropy embodied in folding: whereas the interaction enthalpies favour the folded state, the entropy of the polymer chain favours the unfolded state. Stabilisation from enthalpy gains is offset to a large extent by destabilisation from entropy losses during folding, but these changes are not synchronised, producing barriers in the free energy along the way to the native state (Fig. 1A). These barriers, representing unstable conformations known as transition states, dominate the dynamics of the folding by acting as kinetic bottlenecks in the folding reaction, and they define the folding mechanism via the set of conformations visited during the structural transition.

Energy barriers create a critical challenge for experimental studies of folding, however: being unstable, they are hardly ever occupied. Molecules spend almost all of their time folded or unfolded (or possibly in metastable intermediates), and it has proven extremely difficult to observe them in the middle of the folding transition. Unfortunately, this is the most interesting part of the reaction, since it tells us how the self-assembly happens! Experimentalists have thus been limited to deducing the properties of the transition states indirectly, by observing how changes of experimental conditions (e.g., temperature, pressure, chemical denaturants, mutations, . . .) alter the lifetimes of the visible states and then modeling the effects on the transition states [4].

SINGLE-MOLECULE APPROACHES TO TRANSITION PATHS

Advances in experimental methods in the last few years, in particular methods for observing the folding of single

Michael T. Woodside,

Department of Physics, University of Alberta, Edmonton AB T6G 2E1

National Institute for Nanotechnology, National Research Council, Edmonton AB T6G 2M9

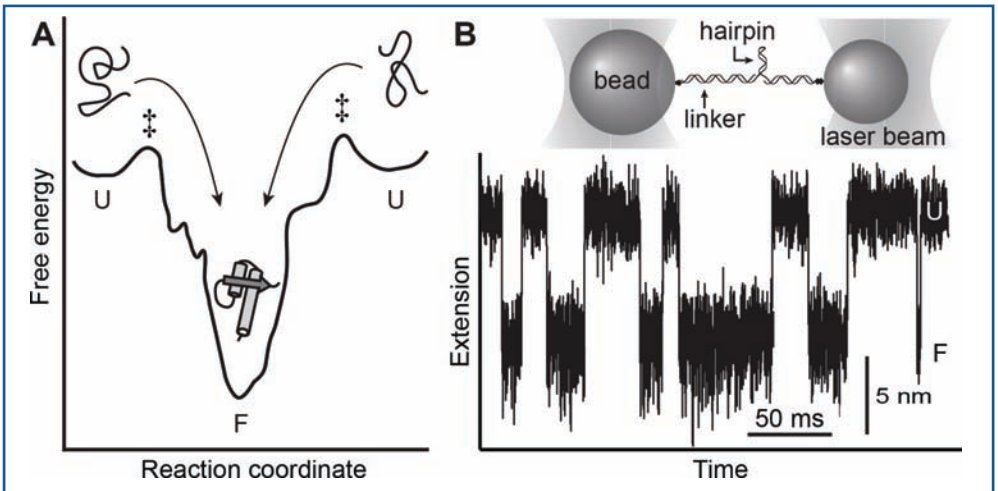

Fig. 1 (A) Cartoon of energy landscape for folding, showing change in free energy from unfolded (U) to folded (F). Landscape is funnel-shaped, but contains barriers (‡) dominating folding kinetics. (B) Folding trajectory of DNA hairpin measured with optical tweezers (adapted from Ref. [11]). Hairpin hops in equilibrium between F (low extension) and U (high extension). Inset: schematic of measurement configuration showing hairpin connected via linkers to beads held in traps.

molecules, have made it possible to characterise the paths followed during folding transitions more directly. Single-molecule approaches are necessary in order to avoid the effects of averaging, because folding transitions cannot be synchronised in an ensemble of molecules. To observe transition paths successfully, however, additional technical hurdles must be overcome: obtaining high enough precision to distinguish different structural states along the paths, high enough time resolution to capture the molecule during the very fleeting transition paths, and a wide enough dynamic range to cover the typical thousand- to million-fold disparity between the transition time and the lifetime of the folded and unfolded states.

Pioneering work using energy transfer between fluorescent dyes attached to different parts of a protein (thereby allowing structural changes during folding to be monitored) showed that the average time for crossing the transition paths could be measured by carefully analysing the statistics of the photons emitted by the dyes [5,6]. These measurements found an average transition time on the order of 2–10 μs, but they could not distinguish individual transition paths, owing to limitations in the dye photophysics. We took a different experimental approach, using laser tweezers to induce and monitor structural changes via single-molecule force spectroscopy (SMFS). Laser tweezers make use of the gradient force induced in small dielectric objects by the electric field gradient in a tightly-focused laser beam to trap micron-sized beads near the focal point—effectively, the tweezers act as a Hookean spring in three dimensions made of light [7]. In SMFS measurements, each end of a molecule is attached via linkers to the beads held by the tweezers (Fig. 1B inset), allowing force to be applied across the molecule. When the molecule unfolds, the unstructured part is stretched out under the applied tension, resulting in a change in the extension of the molecule [8] (Fig. 1B). Laser tweezers provide an excellent probe for transition paths: extensions can be measured with A-scale resolution, time resolution of <10 μs can be achieved by using small beads and stiff traps, and hundreds or thousands of transitions can be observed in a single molecule by applying just enough tension so that the molecule is poised in equilibrium between the folded and unfolded states, hopping back and forth repeatedly. Furthermore, SMFS provides one of the most reliable ways to measure energy landscapes directly, rather than deducing them from modeling [9].

TESTING FOLDING THEORIES WITH TRANSITION-PATH MEASUREMENTS

We recently applied this approach to make the very first measurements of transition paths in proteins and nucleic acids [10,11]. Making use of improvements in instrumental time resolution, we extracted the paths followed through the transition states directly from trajectories of the extension (as in Fig. 1B) by picking out the parts of the trajectory where the molecule moved all the way from unfolded to folded or vice versa. Examples of the transitions observed in a DNA hairpin, a model system for studying folding that consists of a simple stem-loop (Fig. 2A inset), reveal extremely rich and varied behaviour (Fig. 2A): the timescale for individual transition paths can vary by over an order of magnitude, the local velocity along the paths varies greatly, and brief pauses of variable duration can occur at varied locations. The hairpin can even be seen shuttling back and forth across the energy barrier, providing direct evidence of the diffusive nature of the structural search and definitively confirming that folding cannot be described by classic transition-state theory (as once thought), which does not allow for barrier re-crossing. The 1000-fold disparity between the duration of the transition paths and the lifetime of the unfolded and folded states can be seen clearly by comparing Figs. 1B and 2A. To verify that the signals in the transition paths arise from hairpin folding and not instrumental noise, we repeated the measurements with beads connected by handles but no hairpin, finding that the transitions were much faster and less variable (Fig. 2B).

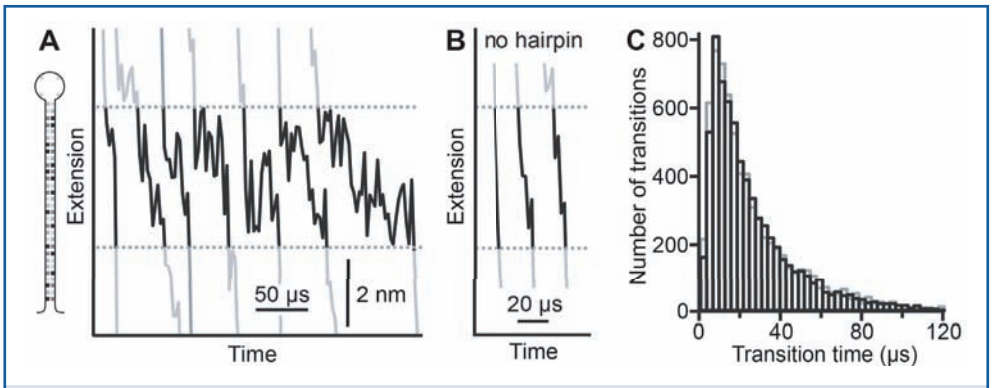

Fig. 2 (A) Representative folding transitions for DNA hairpin showing diverse behaviour. Dotted lines denote boundaries of barrier region. Inset: hairpin structure (dark grey: G:C base-pairs, light grey: A:T base-pairs). (B) 'Transitions' obtained from reference construct without hairpin, illustrating the instrument time response and noise levels. (C) Distribution of transition path times for folding (black) and unfolding (grey). Adapted from Ref. [11].

Measurements like these open an exciting new window onto the biophysics of folding: by allowing the most important parts of folding transitions to be observed directly, they permit the basic physical processes involved to be probed and hence enable experimental testing and refinement of the theory of folding. Furthermore, they hold the promise of providing critical new insights into the mechanisms of folding, by identifying the specific sequence(s) of transition states that determine how a given molecule folds. As an example, we used measurements of the time required to cross the transition paths (Fig. 2C) to test some fundamental features of folding theory. Looking at thousands of transition paths, we found that the average transition time measured directly [11] agreed well with the result predicted for diffusive motions across the energy landscape profile reconstructed experimentally from the equilibrium thermodynamics of the folding [12,13]. Moreover, the coefficient of diffusion for motions along the energy landscape, as extracted from the exponential decay of the transition-time distribution (Fig. 2C), matched the value expected from classic theories of chemical kinetics based on the observed lifetimes of the folded/unfolded states, showing that the theory is consistent with experiment across a wide range of timescales.

We also tested the basic physical picture of folding as a diffusive search over an energy landscape by probing the statistics of the transition-path occupancy—the fraction of time spent at each location within the transition states. Remarkably, although this picture has been tested quantitatively by both computational simulations and experiments combined with modeling, there were no purely experimental tests proving that the observed molecular motions agreed with the expectations for ideal diffusion over the landscape measured for that same molecule. We performed such tests for both model DNA hairpins [14] and the protein that causes "mad-cow" disease, PrP [15], using the transition-path occupancy to determine the conditional probability of being on a transition path at any given value of the molecular extension, $p(\mathrm{TP}|x)$. Theorists predict that for ideal diffusion, $p(\mathrm{TP}|x)$ should be related to the landscape via the splitting probability, $p_{\mathrm{fold}}(x)$, the likelihood that the molecule reaches the folded state before the unfolded state starting from extension x [16]. Intuitively, at the barrier top $p_{\mathrm{fold}} = ½$, whereas $p_{\mathrm{fold}} \sim 1$ near the folded state and ~ 0 near the unfolded state (Fig. 3, inset). Calculating $p(\mathrm{TP}|x)$ from the statistics of the transition paths for these molecules (Fig. 3, black) and comparing to the theoretical expectation for ideal diffusion (Fig. 3, grey) over the measured landscapes (Fig. 3, dashed line) [12,17], we found that indeed there was quantitative agreement for these molecules [14,15].

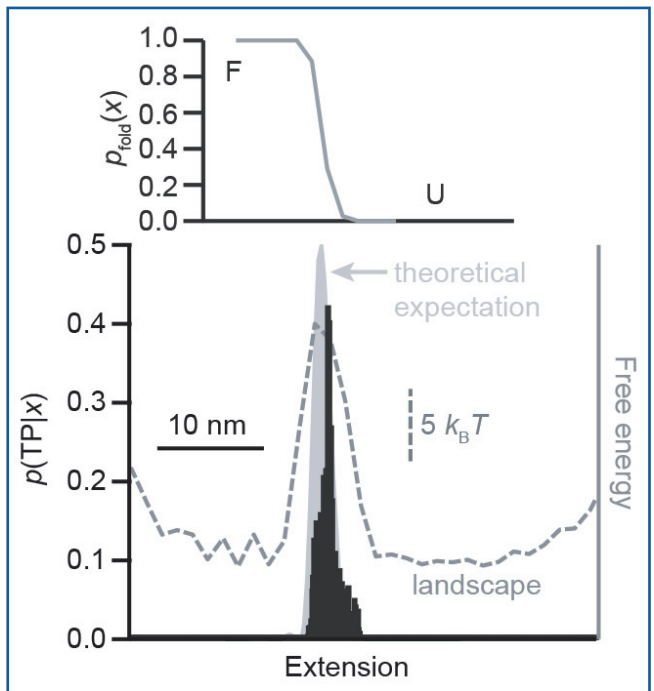

Fig. 3 Conditional transition-path probability (black) for folding of PrP matches theoretical expectation (light grey) calculated from pfold (inset) based on experimentally measured landscape (dashed line), validating basic physical picture of protein folding. F: folded state, U: unfolded state. Adapted from Ref. [15].

FUTURE OUTLOOK

Of course these examples only give a small taste of what can be done with transition-path measurements, and many exciting facets of transition paths remain to be explored. The rich phenomenology of the pauses in the transition paths, which briefly capture the molecule in the transition states, will doubtless reveal fascinating insights into the transition-state properties, and will allow the models heretofore used to characterise transitions states indirectly to be tested directly. Direct observation of transition paths also opens up the ability to see if there are multiple types of pathways and barriers involved in the folding (even if only one dominates) and if so to characterise them individually. Furthermore, opportunities abound for integrating transition-path measurements with long-duration atomistic simulations [18], since the timescales for both are converging. Such integration can amplify the insights that can be obtained from either method alone, both validating the simulations and helping to overcome intrinsic limitations on the structural information available from the experiments. Finally, future work should clarify the effects of the applied force on the transitions, from examining potential kinetic artifacts [19] to testing whether the tension biases the choice of paths. The next few years should thus see many exciting advances in our understanding of folding as transition-path measurements become more widely used.

REFERENCES

1. F. Chiti and C.M. Dobson. "Protein misfolding, functional amyloid, and human disease", *Annu. Rev. Biochem*., **75**, 333-366 (2006).
2. R. Zwanzig, A. Szabo and B. Bagchi, "Levinthal's paradox", *Proc. Natl. Acad. Sci. USA*, **89**, 20-22 (1992).
3. J.N. Onuchic, Z. Luthay-Schulten and P.G. Wolynes, "Theory of protein folding: the energy landscape perspective", *Annu. Rev. Phys. Chem*., **48**, 545-600 (1997).
4. J. Buchner and T. Kiefhaber, eds. Protein Folding Handbook, Wiley, 2006.
5. H.S. Chung, *et al*., "Single-molecule fluorescence experiments determine protein folding transition path times", *Science*, **335**, 981-984 (2012).
6. H.S. Chung and W.A. Eaton, "Single-molecule fluorescence probes dynamics of barrier crossing", *Nature*, **502**, 685-688 (2013).
7. K.C. Neuman and S.M. Block, "Optical trapping", *Rev. Sci. Inst*., **75**, 2787-2809 (2004).
8. D.B. Ritchie and M.T. Woodside, "Probing the structural dynamics of proteins and nucleic acids with optical tweezers", *Curr. Opin. Struct. Biol*., **34**, 43-51 (2015).
9. M.T. Woodside and S.M. Block, "Reconstructing folding energy landscapes by single-molecule force spectroscopy", *Annu. Rev. Biophys*. **43**, 19-39 (2014).
10. H. Yu, *et al*., "Protein misfolding occurs by slow diffusion across multiple barriers in a rough energy landscape", *Proc. Natl. Acad. Sci. USA*, **112**, 8308-8313 (2015).
11. K. Neupane, *et al*., "Direct observation of transition paths during the folding of proteins and nucleic acids", *Science*, **352**, 239-242 (2016).
12. M.T. Woodside, *et al*., "Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid", *Science*, **314**, 1001-1004 (2006).
13. K. Neupane, *et al*., "Transition path times for nucleic acid folding determined from energy-landscape analysis of single-molecule trajectories", *Phys Rev Lett*., **109**, 068102 (2012).
14. K. Neupane, *et al*., "Transition-path probability as a test of reaction-coordinate quality reveals DNA hairpin folding is a one-dimensional diffusive process", *J. Phys. Chem. Lett*., **6**, 1005-1010 (2015).
15. K. Neupane, A.P. Manuel, and M.T. Woodside, "Protein folding trajectories can be described quantitatively by one-dimensional diffusion over measured energy landscapes", *Nature Phys*., **12**, 700-703 (2016).
16. R.B. Best and G. Hummer, "Reaction coordinates and rates from transition paths", *Proc. Natl. Acad. Sci. USA*, **102**, 6732-6737 (2005).
17. Yu, *et al*., "Energy landscape analysis of native folding of the prion protein yields the diffusion constant, transition path time, and rates", *Proc. Natl. Acad. Sci. USA*, **109**, 14452-14457 (2015).
18. H.S. Chung, *et al*., "Structural origin of slow diffusion in protein folding", *Science*, **349**, 1504-1510 (2015).
19. K. Neupane and M.T. Woodside, "Quantifying instrumental artifacts in folding kinetics measured by single-molecule force spectroscopy", *Biophys. J*., **111**, 283-286 (2016).