The capillarity picture and the kinetics of one-dimensional protein folding

Diego U. Ferreiro* and Peter G. Wolynes

Department of Chemistry and Biochemistry and the Center for Theoretical Biological Physics, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0371

Energy landscape theory unites the study of protein folding with the theory of phase transitions. Dimensionality, whose key role in phase transitions is well known, comes to the fore in folding repeat proteins. Repeat proteins are made of near repetitions of 20–40 residues encoding recurring structural motifs. Each repeating element interacts only with its immediate neighbors, forming extended, globally one-dimensional structures that can be interrupted by thermally excited defects (1). The equilibrium properties of repeat proteins may be mapped onto a one-dimensional Ising model (2–4). Going beyond equilibrium, in this issue of PNAS, Werbeck *et al.* (5) delight us by studying the folding kinetics of a very long repeat protein, D34, a 12-ankyrin repeat fragment of AnkyrinR.

Werbeck and Itzhaki (6) previously showed that, unlike short ankyrin repeat proteins, D34 populates an equilibrium folding intermediate at mild denaturing conditions, where the 12-repeat array is neither completely folded nor unfolded. They showed that this "trap" corresponds to an ensemble where the structured regions are polarized toward the C-terminal repeats. Nonuniform stability along the "superhelical" array yields two main cooperative folding "subdomains." As in the classical helix–coil transition of secondary structures, the intrinsic stability of the individual folding elements is low compared with the free energy of stabilization gained by forming an "interface" between neighbors (1). The delicate balances of free energy in each element allows their folding to decouple and subdomains to emerge (7, 8). The present time-resolved experiments show how the one-dimensionality controls the dynamics.

Monitoring the fluorescence, unfolding D34 shows a fast process yielding essentially the equilibrium intermediate. The barrier separating the native from the intermediate states appears "broad," consistent with the picture of neighboring repeats unfolding sequentially (Fig. 1). A slower unfolding phase corresponds to unravelling the C-terminal subdomain. Mutations in this region affect both the rate and its urea dependence in a manner consistent with unfolding by two parallel routes. In principle, parallel routes are expected from the symmetric topology of repeat proteins (8). These have been experimentally traced in the shorter, four-ankyrin-repeat protein Myotrophyn (9). In repeat proteins such distinct populated folding routes are not guaranteed to appear because the routes are selected based on the local energetics, and small perturbations easily reroute the transitions (8–11). For one-dimensional systems the details of the kinetic routes taken through the landscape crucially depend on inhomogeneities in the distribution of energies and entropy losses for folding along the array.

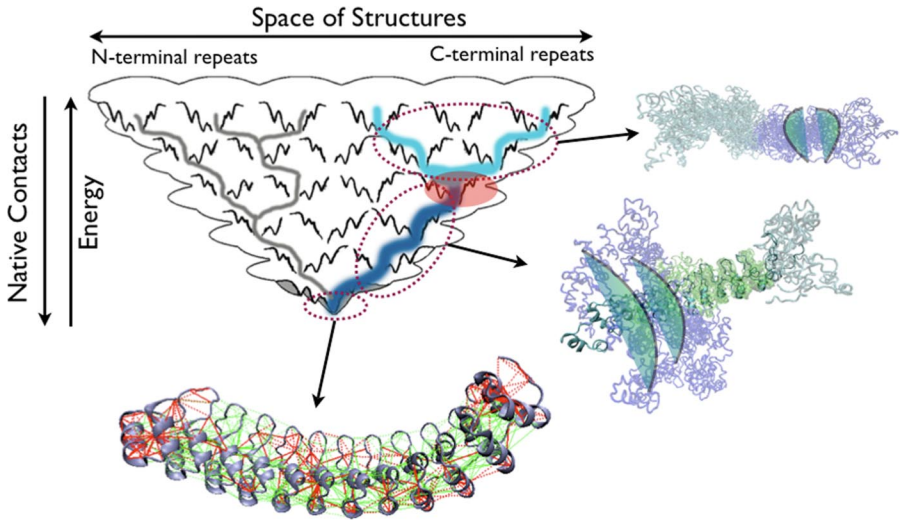

Fig. 1. Navigating the repeat-protein landscape. Each repeat has similar structural elements that interact with nearest neighbors, depicted by many folding funnels that merge on interaction and comprise one overall funneled landscape (11). High-energy unfolded configurations are near the top and the native state is at the bottom. In between, an intermediate ensemble of partially folded states becomes populated at equilibrium (red shadow). Inhomogeneities in the local energies and entropy costs for folding mold the landscape, defining the actual routes. The thick blue lines depict the preferred unfolding routes of the 12-ankyrin repeat protein D34 (5). The thin gray lines depict less frequent routes that may account for spectroscopically silent phases. The native structure of D34 is pictured at the bottom, colored according to its local frustration pattern (18). Minimally frustrated interactions are pictured as green, whereas highly frustrated interactions are shown in red. On the right are ensembles of structures along the folding transitions that were obtained from molecular dynamic calculations. The black lines represent a "front" that separates the folded from the unfolded regions.

Because intramolecular forces are short-ranged, repeat proteins fold through configurations in which residues contiguous in physical space will be either ordered or disordered, like droplets in a phase transition. The capillarity picture of folding (12) thus may be used to describe folding in a large protein. In the capillarity model, a front between the folded and unfolded parts can be defined. For elongated architectures this front orients orthogonal to the long axis and moves in a one-dimensional fashion. The nucleation of the capillarity front requires overcoming a free-energy barrier. The rate of the front's propagation is reflected in a prefactor depending on the local landscape ruggedness and the friction from the solvent (12). The motion of the front is impeded by transient trapping in local minima, and when the residence time becomes too long, the traps can be thought as minifunnels within the main funnel (Fig. 1). Entropy is gained in stages and unfolding becomes intermittent (13), as often seen in mechanical stretching experiments (14). The measured rate of folding ($1/\tau$) de-

Author contributions: D.U.F. and P.G.W. wrote the paper.

The authors declare no conflict of interest.

See companion article on page 9982.

*To whom correspondence should be addressed. E-mail: dferreiro@ucsd.edu.

pends on the rate of moving the front and the barrier height for forming it (ΔG): $\tau = 2\pi\tau_{corr}e^{\Delta G/k_BT}$. In the absence of internal friction, τ_{corr} is approximately the Eaton speed limit, ≈100 ns (15). By using this prefactor, the rates measured for D34 yield a kinetic defect formation barrier of ≈6.7 kcal·mol^{-1} for the first refolding transition. This kinetic barrier should be close (but not precisely equal) to the equilibrium free-energy cost of introducing a defect determined by the purely thermodynamic experiments. Studies on the 7AR domain of Notch give an estimated stability for single AR repeat of ≈6.5 kcal·mol^{-1}, whereas forming the joining interface yields −9 kcal·mol^{-1} (4). Using similar values for D34 suggests that the main barrier corresponds to the folding of approximately two or three consecutive repeats.

The kinetics of D34 in fact reveals a kinetic trap. Energy landscape theory suggests such a trap may arise purely topologically (8) or as an effects of "frustration" (16). When frustration is particularly low, the strong energetic bias toward the native basin overcomes the small asperities of the landscape and ultimately plays off the dominant topology-dependent entropy of the chain (17). The one-dimensionality of repeat proteins makes the effect of local frustration larger than in three-dimensional globular proteins, much as a traffic jam arises on a crowded (one dimensional!) freeway. Such local frustration might be at work in D34. Theoretical methods to spatially localize and quantify the energetic frustration present in native protein structures (18) reveal three regions of highly frustrated interactions in D34 in an otherwise strongly cross-linked web of minimally frustrated interactions (bottom of Fig. 1). The highly frustrated regions are located at the ends, and in a central region between repeats 5 and 6. These breaks occur precisely where the experiments suggest the intermediate's folding boundary is. It is interesting to speculate that the highly frustrated interactions are responsible for the roughness that gives rise to a long-lived folding intermediate. Modification of this interface so that it is minimally frustrated should destabilize the trap, making the D34 landscape appear much smoother.

The capillarity picture explains several puzzling features of repeat protein folding kinetics.

The capillarity picture explains several puzzling features of repeat protein folding kinetics. Natural ankyrin repeat proteins fold 4 orders of magnitude slower than expected from their contact order, whereas other repeat proteins with similar contact order fold much faster (1). The capillarity picture suggests that the differences in rates are related to the escape from local traps. Recent experiments on redesigned repeat proteins confirm this (19). When the local frustration is minimized by using "consensus" repeats, the kinetic barrier appears to be considerably lower (4.5 kcal·mol^{-1} for Ankyrin, 2.3 kcal·mol^{-1} for TPR).

The modular nature of repeat proteins gives a highly versatile framework for evolving novel binding properties (20). The collective influence of local interactions along the one-dimensional scaffold allows sites to thermodynamically modulate each other even at a considerable distance (2). The folding of repeat proteins is strongly coupled to their biological behavior, in particular, to their degradation (10, 21). The one-dimensional view of repeat proteins will help us to understand how protein function emerges from the folding landscape, a challenge that can be called the "second part" of the folding problem.

ACKNOWLEDGMENTS. D.U.F. is a Jane Coffin Childs Fellow. This work was supported in part by National Institutes of Health Grant PO1GM071862.

1. Kloss E, Courtemanche N, Barrick D (2008) Repeat-protein folding: New insights into origins of cooperativity, stability, and topology. *Arch Biochem Biophys* 469:83–99.
2. Ferreiro DU, Walczak AM, Komives EA, Wolynes PG (May 16, 2008) The energy landscapes of repeat-containing proteins: Topology, cooperativity, and the folding funnels of one-dimensional architectures. *PLoS Comput Biol* 4: e1000070.
3. Kajander T, Cortajarena AL, Main ER, Mochrie SG, Regan L (2005) A new folding paradigm for repeat proteins. *J Am Chem Soc* 127:10188–10190.
4. Mello CC, Barrick D (2004) An experimentally determined protein folding energy landscape. *Proc Natl Acad Sci USA* 101:14102–14107.
5. Werbeck ND, Rowling PJE, Ranjani V, Itzhaki LS (2008) Shifting transition states in the unfolding of a large ankyrin repeat protein. *Proc Natl Acad Sci USA* 105:9982–9987.
6. Werbeck ND, Itzhaki LS (2007) Probing a moving target with a plastic unfolding intermediate of an ankyrin-repeat protein. *Proc Natl Acad Sci USA* 104:7863–7868.
7. Bradley CM, Barrick D (2002) Limits of cooperativity in a structurally modular protein: Response of the Notch ankyrin domain to analogous alanine substitutions in each repeat. *J Mol Biol* 324:373–386.
8. Ferreiro DU, Cho SS, Komives EA, Wolynes PG (2005) The energy landscape of modular repeat proteins: Topology determines folding mechanism in the ankyrin family. *J Mol Biol* 354:679–692.
9. Lowe AR, Itzhaki LS (2007) Rational redesign of the folding pathway of a modular protein. *Proc Natl Acad Sci USA* 104:2679–2684.
10. Barrick D, Ferreiro DU, Komives EA (2008) Folding landscapes of ankyrin repeat proteins: Experiments meet theory. *Curr Opin Struct Biol* 18:27–34.
11. Ferreiro DU, Komives EA (2007) The plastic landscape of repeat proteins. *Proc Natl Acad Sci USA* 104:7735–7736.
12. Wolynes PG (1997) Folding funnels and energy landscapes of larger proteins within the capillarity approximation. *Proc Natl Acad Sci USA* 94:6170–6175.
13. Socci ND, Onuchic JN, Wolynes PG (1996) Diffusive dynamics of the reaction coordinate for protein folding funnels. *J Chem Phys* 104:5860–5868.
14. Li L, Wetzel S, Pluckthun A, Fernandez JM (2006) Stepwise unfolding of ankyrin repeats in a single protein revealed by atomic force microscopy. *Biophys J* 90:L30–L32.
15. Kubelka J, Hofrichter J, Eaton WA (2004) The protein folding 'speed limit.' *Curr Opin Struct Biol* 14:76–88.
16. Bryngelson JD, Wolynes PG (1987) Spin glasses and the statistical mechanics of protein folding. *Proc Natl Acad Sci USA* 84:7524–7528.
17. Clementi C (2008) Coarse-grained models of protein folding: Toy models or predictive tools? *Curr Opin Struct Biol* 18:10–15.
18. Ferreiro DU, Hegler JA, Komives EA, Wolynes PG (2007) Localizing frustration in native proteins and protein assemblies. *Proc Natl Acad Sci USA* 104:19819–19824.
19. Wetzel SK, Settanni G, Kenig M, Binz HK, Pluckthun A (2008) Folding and unfolding mechanism of highly stable full-consensus ankyrin repeat proteins. *J Mol Biol* 376:241–257.
20. Bjorklund AK, Ekman D, Elofsson A (2006) Expansion of protein domain repeats. *PLoS Comput Biol* 2:e114.
21. Truhlar SM, Mathes E, Cervantes CF, Ghosh G, Komives EA (2008) Pre-folding IkappaBalpha alters control of NF-kappaB signaling. *J Mol Biol* 380:67–82.