Chemical physics of protein folding

Peter G. Wolynes[a,1], William A. Eaton[b], and Alan R. Fersht[c]

[a]*Department of Chemistry and Center for Theoretical Biological Physics, Rice University, Houston, TX 77005;* [b]*National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892; and* [c]*Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 0QH, United Kingdom*

In 2008, to the consternation of some, one of the editors of this special issue on the "Chemical Physics of Protein Folding," was quoted as saying, "What was called the protein folding problem 20 years ago is solved" (1). One purpose of this special issue is to drive home this point. The other, more important purpose is to illustrate how workers on the protein folding problem, by moving beyond their early obsession with seeming paradoxes (2), are developing a quantitative understanding of how the simpler biological structures assemble both in vitro and in vivo. The emerging quantitative understanding reveals simultaneously the richness of folding phenomena and the elegant simplicity of the underlying principles of spontaneous biomolecular assembly. The appreciation of these contrasting aspects of the folding problem has come about through the cooperation of theorists and experimentalists, a theme common to all the contributions to this special issue. Although the basic ideas about the folding energy landscape have turned out to be quite simple, entering even into some undergraduate textbooks (3), exploring their consequences in real systems has required painstaking intellectual analysis, as well as detailed computer simulations and experiments that still stretch the bounds of what is feasible. The backgrounds of the contributors to this issue reflect the breadth of the folding field and range from computer science and theoretical physics to molecular biology and organic chemistry. A great deal of the progress in the field can thus be traced to a fairly successful effort to develop a common language and conceptual framework for describing folding.

The conceptual framework is provided by energy landscape theory, which describes the diversity of structural possibilities in statistical mechanical terms. The main paradoxes of folding are resolved by the consistency principle (4) or, more generally, by the principle of minimal frustration (5), which quantifies the dominance of interactions stabilizing the specific native structure over other interactions that would favor nonnative, topologically distinct traps. In other words, the energy landscape of evolved proteins appears to be funneled (6, 7).

The overall funneled nature of the folding landscape provides a first guess of how folding begins and continues: Proteins fold by assembling primarily native substructures, whereas they only transiently sample misfolds. This insight explains the success of protein engineering in providing detailed structural information on the transition state ensemble, the so-called "φ-value analysis" (8, 9). In some cases, changing solution conditions changes the position of the transition state along the reaction coordinate, such that φ-value studies can even tell us in what order native parts assemble along the dominant transition path.

The ensemble nature of the transition states was the first clue that mechanistic complexity still remains on funneled landscapes. Multiple choices of a precise folding mechanism are possible, and several experimental studies herein demonstrate the malleability of specific folding mechanisms for proteins of either related structure or related sequence; on a funneled landscape, many roads lead to Rome (10, 11). After a molecule embarks on a route, even on a funneled energy landscape, assembly may not be completely straightforward. Occasionally, a greedy attempt to make native contacts early on can lead to topological traps (12); in that case, some early native interactions must be undone to allow complete folding and one must backtrack (13). Also, evolution toward a funneled landscape cannot repeal the universal character of the physics of specific molecular interactions; thus, no real folding landscape is ever perfectly funneled. If the nonnative interactions are fairly weak, they just provide a source of "friction," a topic quantified here in several papers (14, 15). If the nonnative interactions become stronger, as theory predicts, frustrated interactions can allow specific intermediates to form with substantial nonnative contacts along with some native structure. This phenomenon also receives attention here (16).

Many of the most intriguing questions of mechanism can only be addressed feebly at the level of ensemble-averaged experiments; thus, folding science has called forth some extremely challenging experiments in which molecules are studied individually one at a time (17–19). On the computational side, although effective theories of folding can often use simplified models that make interesting and surprising predictions (20, 21), many questions of detail can now be answered satisfactorily with fully atomistic simulations that challenge computational power to the limit (22–24).

As the title of this special issue suggests, most of the contributions focus on exploring the general principles of folding. These principles emerge as clearly in test tube studies as they do in the cell. However, the cell provides its own challenges to folding science. The effects of heterogeneous intracellular environments on folding landscapes (25) are beginning to be explored. Also, despite the inborn tendency of proteins to do the right thing (arising from the minimal frustration principle), alternate folding, misfolding, and aggregation for a few proteins do lead to pathology, a topic explored in this issue (26) with the superoxide dismutase system that is involved in amyotrophic lateral sclerosis.

The topics discussed in this issue are only a small part of the work in the folding field. Nevertheless, they make clear that protein folding is a vibrant, living, interdisciplinary part of the natural sciences. We hope this snapshot of the field will encourage others to bring new approaches to contribute to our understanding of biomolecular assembly and encourage the use of the ideas and strategies that have already proved successful in the study of assembling the simplest biomolecules to look at the full complexity of living systems at higher levels of complexity.

Author contributions: P.G.W., W.A.E., and A.R.F. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: pwolynes@rice.edu.

1. Service RF (2008) Problem solved* (*sort of). *Science* 321(5890):784–786.
2. Wolynes PG, Eaton WA (1999) The physics of protein folding. *Physics World* 12(9):39–44.
3. Voet D, Voet JG, Pratt CW (1999) *Fundamentals of Biochemistry* (Wiley, New York), p 154.
4. Go N (1983) Theoretical studies of protein folding. *Annu Rev Biophys Bioeng* 12:183–210.
5. Bryngelson JD, Wolynes PG (1987) Spin glasses and the statistical mechanics of protein folding. *Proc Natl Acad Sci USA* 84(21):7524–7528.
6. Leopold PE, Montal M, Onuchic JN (1992) Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proc Natl Acad Sci USA* 89(18):8721–8725.
7. Kubelka J, Henry ER, Cellmer T, Hofrichter J, Eaton WA (2008) Chemical, physical, and theoretical kinetics of an ultrafast folding protein. *Proc Natl Acad Sci USA* 105 (48):18655–18662.
8. Fersht AR (1999) *Structure and Mechanism in Protein Science* (Freeman, New York).
9. Matouschek A, Kellis JT, Jr., Serrano L, Fersht AR (1989) Mapping the transition state and pathway of protein folding by protein engineering. *Nature* 340 (6229):122–126.

www.pnas.org/cgi/doi/10.1073/pnas.1215733109

10. Giri R, et al. (2012) Folding pathways of proteins with increasing degree of sequence identities but different structure and function. *Proc Natl Acad Sci USA* 109:17772–17776.
11. Korzhnev DM, Religa TL, Kay LE (2012) Transiently populated intermediate functions as a branching point of the FF domain folding pathway. *Proc Natl Acad Sci USA* 109:17777–17782.
12. Sułkowska JI, Noel JK, Onuchic JN (2012) Energy landscape of knotted protein folding. *Proc Natl Acad Sci USA* 109:17783–17788.
13. Li W, Terakawa T, Wang W, Takada S (2012) Energy landscape and multiroute folding of topologically complex proteins adenylate kinase and 2out-knot. *Proc Natl Acad Sci USA* 109:17789–17794.
14. Wensley BG, et al. (2012) Separating the effects of internal friction and transition state energy to explain the slow, frustrated folding of spectrin domains. *Proc Natl Acad Sci USA* 109:17795–17799.
15. Soranno A, et al. (2012) Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proc Natl Acad Sci USA* 109:17800–17806.
16. Beauchamp KA, McGibbon R, Lin Y-S, Pande VS (2012) Simple few-state models reveal hidden complexity in protein folding. *Proc Natl Acad Sci USA* 109:17807–17813.
17. Stigler J, Rief M (2012) Calcium-dependent folding of single calmodulin molecules. *Proc Natl Acad Sci USA* 109:17814–17819.
18. Jagannathan B, Elms PJ, Bustamante C, Marqusee S (2012) Direct observation of a force-induced switch in the anisotropic mechanical unfolding pathway of a protein. *Proc Natl Acad Sci USA* 109:17820–17825.
19. Ferreon ACM, Moosa MM, Gambin Y, Deniz AA (2012) Counteracting chemical chaperone effects on the single-molecule α-synuclein structural landscape. *Proc Natl Acad Sci USA* 109:17826–17831.
20. Reddy G, Liu Z, Thirumalai D (2012) Denaturant-dependent folding GFP. *Proc Natl Acad Sci USA* 109:17832–17838.
21. Shental-Bechor D, et al. (2012) Nonnative interactions regulate folding and switching of myristoylated protein. *Proc Natl Acad Sci USA* 109:17839–17844.
22. Piana S, Lindorff-Larsen K, Shaw DE (2012) Protein folding kinetics and thermodynamics from atomistic simulation. *Proc Natl Acad Sci USA* 109:17845–17850.
23. McCully ME, Beck DAC, Daggett V (2012) Multimolecule test-tube simulations of protein unfolding and aggregation. *Proc Natl Acad Sci USA* 109:17851–17856.
24. Potoyan DA, Papoian GA (2012) Regulation of the H4 tail binding and folding landscapes via Lys-16 acetylation. *Proc Natl Acad Sci USA* 109:17857–17862.
25. Guo M, Xu Y, Gruebele M (2012) Temperature dependence of protein folding kinetics in living cells. *Proc Natl Acad Sci USA* 109:17863–17867.
26. Lang L, Kurnik M, Danielsson J, Oliveberg M (2012) Fibrillation precursor of superoxide dismutase 1 revealed by gradual tuning of the protein-folding equilibrium. *Proc Natl Acad Sci USA* 109:17868–17873.