PERSPECTIVES: STATISTICAL THERMODYNAMICS

Taking a Walk on a Landscape

Charles L. Brooks III, José N. Onuchic, David J. Wales

The folding of proteins, the complex behavior of glasses, and the structure and dynamics of atomic and molecular clusters have long been studied in separate disciplines. In recent years, energy landscape theory has emerged as a unifying language for experimentalists and theorists to describe structure formation and dynamics in these complex systems. Researchers in all three disciplines met to exchange ideas at a symposium organized as part of the American Chemical Society (ACS) National Meeting in San Diego, California, in April.

The energy landscape describes how the energy of a system changes with geometry, as defined by the coordinates of atoms, molecules, or side chains. At a minimum, a small displacement in any direction increases the potential energy, just as in a basin surrounded by mountains, a step in any direction is uphill. Potential energy surfaces of complex systems usually have vast numbers of local minima; the lowest one—the deepest basin—is the global minimum. Energy landscapes for different systems may differ widely. These differences are responsible for the fact that natural proteins and crystals can reliably locate one particular structure from many possible ones, whereas glasses fail to do so.

Levinthal established the classical viewpoint of protein folding by noting that proteins cannot fold by a random search for the native (correctly folded) state because this would take far too long, and must therefore fold via a sequential, energy biased folding pathway. An energy bias favoring the native state allows proteins to fold within a biologically relevant time frame. For proteins with random amino acid sequences, one would expect no such bias and a rough potential energy surface, reflecting the impossibility of incorporating all favorable interactions into the same structure: The system is said to be frustrated. Natural proteins have presumably evolved to enhance this native state bias and to minimize frustration.

C. L. Brooks III is in the Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA. E-mail: brooks@scripps.edu J. N. Onuchic is in the Department of Physics, University of California at San Diego, La Jolla, CA 92093, USA. E-mail: jonuchic@ucsd.edu D. J. Wales is in the Department of Chemistry, Cambridge University, Lensfield Road, Cambridge CB2 1EW, UK. E-mail: dw34@cam.ac.uk

Generalizing and replacing the classical view of folding as a single stepwise pathway with a series of intermediates, the energy landscape theory of protein folding (*1–3*) involves the progressive organization of an ensemble of partially folded structures on a free energy surface that resembles a partially rough funnel (see the first figure). Local free energy minima correspond to transient kinetic traps. The native state lies at the bottom of a set of convergent pathways.

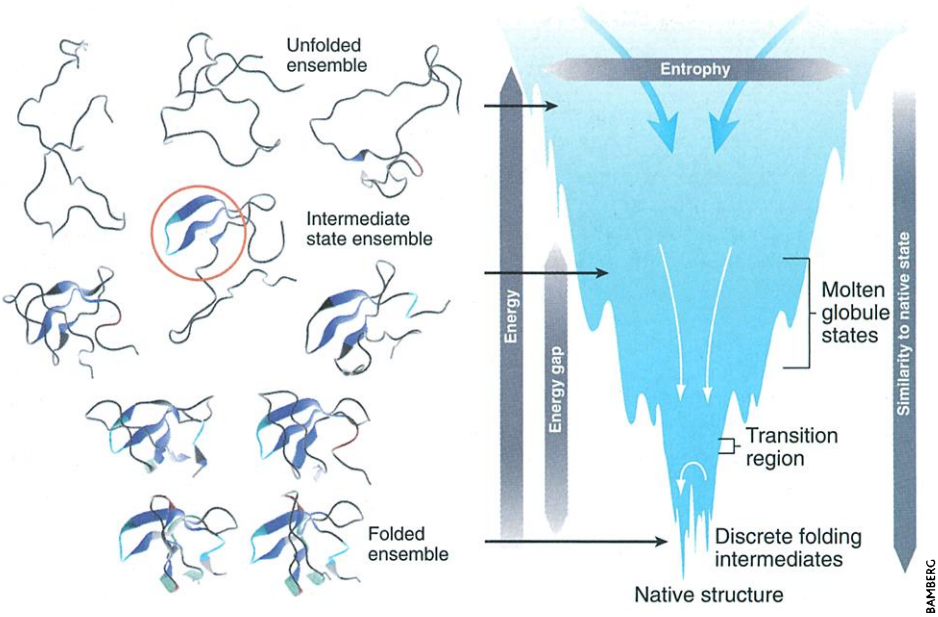
The key role of the energy surface in determining the properties of glasses was recognized by Goldstein around the time that Anfinsen and Levinthal made seminal contributions to protein science. Substantial progress in calculating thermodynamic properties directly from knowledge of local minima was made in the 1980s (*4*). This approach has been extended to dynamical properties and applied to a wide variety of atomic and molecular clusters (*5*, *6*). The properties of different types of glasses have also been interpreted in terms of global features of the energy landscape (*7*). More recently, it has been suggested that glassy liquids are composed of a mosaic of clusters, each of which has its own landscape (*8*).

The energy landscape for each of these systems—proteins, glasses, and clusters—can be conceptualized as a low-dimensional free energy surface, with two axes depicting the energy-entropy balance associated with competing physical processes and a ruggedness denoting the presence of frustration that leads to specific kinetic behavior (see the first figure).

Potential energy surfaces can be visualized more quantitatively with a "disconnectivity graph" (*9*). Such graphs are constructed from databases of stationary points on the many-dimensional potential energy surface and the pathways connecting them (see A and B in the second figure). The graphs divide distinct local minima into separate branches that can only interconvert if a threshold energy is exceeded. A "palm-tree" disconnectivity graph corresponds to a free energy landscape with funneling characteristics.

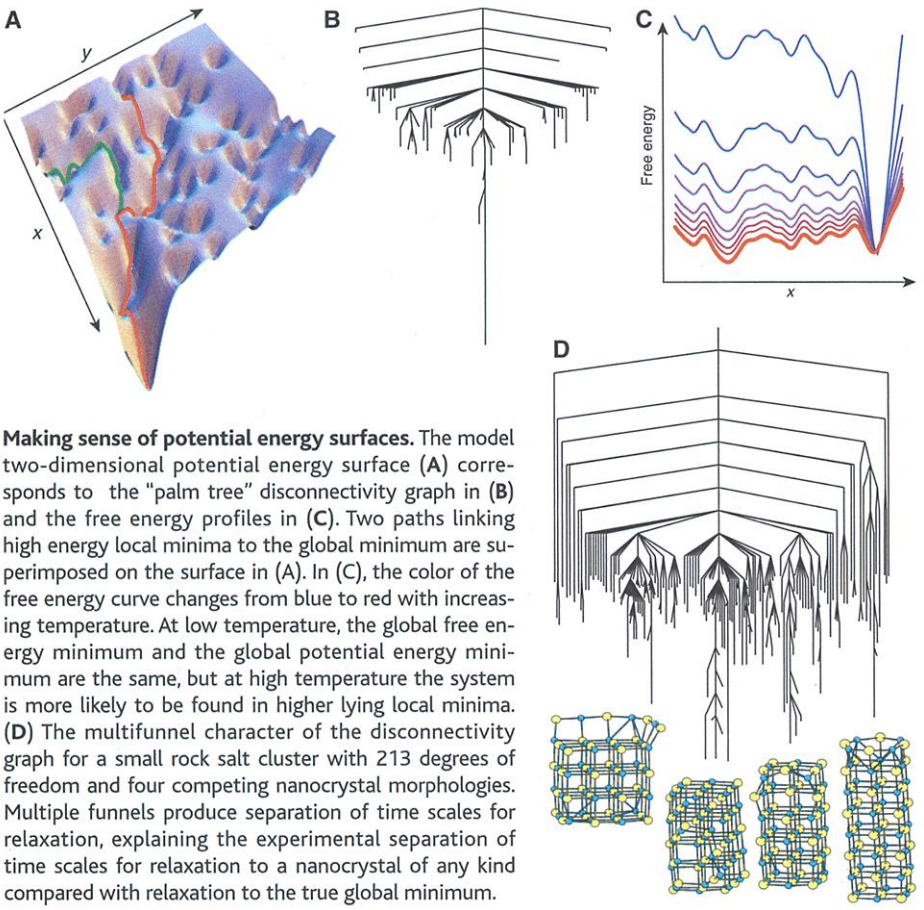
Direct dynamic calculations based on such graphs have been carried out for peptides and numerous clusters (*6*). These calculations reach the experimental time scale but have not yet been extended to proteins with more than 20 residues. In the future, such calculations may help to describe the properties of bulk glasses and proteins at the level of detail achieved to date only with smaller systems, such as the alkali halide clusters that are being explored with ion chromatography (*10*) (see D in the second figure).

For proteins, the potential energy landscape can be connected to a free energy



The free energy landscape for protein folding. Folding occurs through the progressive organization of ensembles of structures [shown here for the src-SH3 domain (**left**)] on a funnel-shaped free energy landscape (**right**). Conformational entropy loss during folding is compensated by the free energy gained as more native interactions are formed. Kinetics is determined by the local roughness of the landscape, relative to thermal energy. Key interactions in early folding (dashed circle) coincide, for this protein, with experimentally determined regions.

CREDIT: C. CLEMENTI AND G. BAMBERG



Making sense of potential energy surfaces. The model two-dimensional potential energy surface (**A**) corresponds to the "palm tree" disconnectivity graph in (**B**) and the free energy profiles in (**C**). Two paths linking high energy local minima to the global minimum are superimposed on the surface in (A). In (C), the color of the free energy curve changes from blue to red with increasing temperature. At low temperature, the global free energy minimum and the global potential energy minimum are the same, but at high temperature the system is more likely to be found in higher lying local minima. (**D**) The multifunnel character of the disconnectivity graph for a small rock salt cluster with 213 degrees of freedom and four competing nanocrystal morphologies. Multiple funnels produce separation of time scales for relaxation, explaining the experimental separation of time scales for relaxation to a nanocrystal of any kind compared with relaxation to the true global minimum.

landscape by projecting the energy surface onto one or two important coordinates, such as the number of native contacts and the radius of gyration. Free energy surfaces, like disconnectivity trees, can be used to calculate thermodynamics and kinetics for problems such as protein folding.

Computer simulation is used to provide quantitative predictions for all the systems considered here. The most detailed views of protein folding come from fully atomistic simulation models, which have been used to explore folding free energy surfaces and folding mechanisms (*3*, *11*), protein unfolding and correlation with measured data on the folding transition state (*12*), and direct folding of small polypeptides (*13*). Simulations with explicit solvent models have shown how water "lubricates" the search for native interactions (*3*).

Simulations with simplified representations of proteins have shown how topology and energy interact to yield specific folding scenarios for individual proteins and helped to interpret experimental results pertaining to the folding transition state (*1*, *14*). These models are used to interpret experimental folding kinetics (*12*) and explore the role of side chain entropy in folding (*15*, *16*). Related analytical models are also being developed for protein design.

Detailed connections between dynamical simulations and the underlying potential energy surface for model glasses have also started to appear (*17*, *18*). These calculations have identified a "landscape-dominated" regime, where the barriers between local minima start to increase and structural relaxation is slowed down with decreasing temperatures (*18*).

Experimental progress is not lagging behind. Early folding events can now be monitored on nanosecond to microsecond time scales, enabling details of the landscape near the top of the folding funnel to be probed (*19*, *20*). Peptides with well-defined structures have been designed, and are used to determine the formation time for basic folding structures such as α-helices and β-hairpins (*21*). The response of folding kinetics to mutations has been used to determine the participation of different parts of the protein in the folding transition state ensemble and, for larger proteins, in the intermediate ensembles (*12*).

The energy landscape approach has also been used to develop new algorithms for global optimization and protein structure prediction (*22*). The principle that the energy landscape of a protein should be funnel-like and minimally frustrated has been used to design knowledge-based potentials for fold recognition (*22–24*).

Related to structure-prediction efforts, combinatorial experiments, aided by theoretical tools, are starting to provide new ways to probe the determinants of protein folding and to identify novel folding amino acid sequences (*25*). Computational and theoretical models that examine the "evolutionary" free energy landscapes of proteins to understand the interrelationships among protein families and the development of drug resistance (*26*) also exploit folding landscape ideas. Finally, the interplay between folding and aggregation and folding and function (*2*, *27*) can be addressed using this ensemble description.

The meeting in San Diego provided a view of a future in which traditional disciplinary boundaries are dissolved and ideas from many fields are integrated to address key challenges of complex systems in biology, chemistry, and physics. Scientists across these areas of investigation are focusing on how experimental observables are determined by potential and free energy landscapes. For proteins, future work must go beyond folding to understand function and critical issues related to diseases caused by misfolded proteins. For glasses, a detailed understanding of how characteristic properties arise from the underlying potential energy surface is required. More accurate calculations to bridge the time scale gap between experiment and simulation are likely to play a key role in each case.

References

1. C. Clementi *et al.*, *J. Mol. Biol.* **298**, 937 (2000).
2. C. M. Dobson *et al.*, *Angew. Chem. Int. Ed.* **37**, 868 (1998).
3. J.-E. Shea, C. L. Brooks III, *Annu. Rev. Phys. Chem.* **52**, 499 (2001).
4. F. H. Stillinger, *Science* **267**, 1935 (1995).
5. R. S. Berry *et al.*, *NATO ASI Ser. B* **158**, 185 (1987).
6. D. J. Wales *et al.*, *Adv. Chem. Phys.* **115**, 1 (2000).
7. C. A. Angell, *Science* **267**, 1924 (1995).
8. X. Y. Xia, P. G. Wolynes, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2990 (2000)
9. O. M. Becker, M. Karplus, *J. Chem. Phys.* **106**, 1495 (1997).
10. P. Dugourd *et al.*, *Chem. Phys. Lett.* **267**, 186 (1997).
11. G. Hummer *et al.*, *Proteins Struct. Funct. Genet.* **42**, 77 (2001).
12. K.-B. Wong *et al.*, *J. Mol. Biol.* **296**, 1257 (2000).
13. M. Shirts, V. S. Pande, *Science* **290**, 1903 (2000).
14. J.-E. Shea *et al.*, *J. Chem. Phys.* **113**, 7663 (2000).
15. J. Shimada *et al.*, *J. Mol. Biol.* **308**, 79 (2001).
16. D. K. Klimov, D. Thirumalai, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2544 (2000).
17. F. Sciortino *et al.*, *J. Phys. Condens. Matter* **12**, 6525 (2000).
18. S. Sastry *et al.*, *Nature* **393**, 554 (1998).
19. M. Gruebele, *Annu. Rev. Phys. Chem.* **50**, 485 (1999).
20. W. A. Eaton *et al.*, *Annu. Rev. Biophys. Biomol. Struct.* **29**, 327 (2000).
21. E. Lacroix *et al.*, *Curr. Opin. Struct. Biol.* **9**, 487 (1999).
22. J. Pillardy *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2329 (2001).
23. J. Skolnick *et al.*, *Nature Biotechnol.* **18**, 283 (2000).
24. C. Hardin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14235 (2000).
25. H. Kono, J. G. Saven, *J. Mol. Biol.* **306**, 607 (2001).
26. J. D. Hirst, *Protein Eng.* **12**, 721 (1999).
27. J. M. Finke *et al.*, *Biochemistry* **39**, 15633 (2000).