Published in final edited form as:

Curr Opin Struct Biol. 2011 February ; 21(1): 1–3. doi:10.1016/j.sbi.2010.12.003.

Protein Folding and Binding: From Biology to Physics and Back Again

Michael K. Gilson and

Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0736

Sheena E. Radford

Astbury Centre for Structural Molecular Biology and the Institute of Molecular and Cellular Biology, University of Leeds, Leeds, LS2 9JT, UK

Michael K. Gilson: mgilson@ucsd.edu; Sheena E. Radford: s.e.radford@leeds.ac.uk

Anyone who has seen vivid animations of the molecules of life in action (www.nytimes.com/2010/11/16/science/16animate.html) can scarcely escape a sense that the molecules themselves are alive, as they take on form, find partners, adjust and adapt. Perhaps this impression stems partly from artistic or didactic liberties, as some atomistic jiggling, wiggling and molecular missteps might have been glossed over. Nonetheless, it is entirely reasonable to expect that “life-like” processes emerge from “physics-like” processes at some subcellular level of organization. There is something fascinating about this crossing from physics into biology, and there remain many fundamental issues that are not yet understood. How does life emerge from random thermal motions? We trust that life processes can be accounted for by statistical and quantum mechanics, but physics only establishes a framework for which mechanisms might work. It is the vagaries of evolutionary history that determined what actually came to be.

And there is no guarantee that the actual, microscopical mechanisms of biomolecules conform to our intuitions and expectations, informed as they are by the macroscopic world of human-engineered devices and machines, and by the reductionist appeal of purified systems. Ideas like locks, keys, switches, springs and pathways can be helpful; but they can also mislead in the molecular realm of thermal fluctuations and conformational spaces. In cells, proteins may fold via an ensemble of pathways different from that of isolated molecules folding in the dilute solutions commonly used by biophysicists. Moreover, protein engineering is still quite different from civil or electrical engineering, as we still have not identified molecular components with sufficiently reliable specifications, analogous to girders and transistors. Thus, although much has been learned, we are still enumerating principles of biomolecular mechanism and seeking key variables and useful simplifications. The reviews in this section on Folding and Binding bring a valuable update from the forefront of research into the actual physical mechanisms of biomolecular formation and function, drawing on approaches that span simulation, biophysical methods and cellular investigation.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Bowman *et al.* open the discussion with atomistic insights into the folding of small proteins in isolation, based on a sophisticated new computational approach, Markov state modeling. This yields maximal information from multiple parallel molecular dynamics calculations by grouping trajectory snapshots into clusters of related conformations with well-defined interconversion kinetics. Results to date highlight the complexity of the folding process, even for proteins with apparently simple two-state experimental folding kinetics; suggest that non-native contacts can speed folding to the native state; and point to a rather unexpected pattern of conformational transitions in which different non-native conformations interconvert more rapidly via the native conformation than via their non-native counterparts.

Sosnick and Barrick then provide an experimental perspective of the folding mechanisms of simple proteins in isolation, highlighting persistent uncertainties and controversies in this now well matured field. These authors debate the relevance of residual structure in the denatured states of proteins in prefiguring the folded conformation and discuss whether such interactions can speed folding once denaturant is removed. They also review the effects of removal of denaturant in the early stages of folding (biophysically speaking), debate whether this may lead to collapse of the chain in a generic sense, and consider how this may influence folding. Sosnick and Barrick also address the concept, discussed in the prior article, that non-native states may interconvert chiefly via the native state.

Zhang and Ignatova then move us from the folding of isolated proteins in beakers to folding in a biological context. It is now known that folding can occur during the very process of translation, to the point that alpha-helix forms while the nascent peptide is still within the ribosomal exit tunnel. The potential for the process of co-translational folding to be profoundly different from folding of a complete protein in a test-tube is emphasized by the fact that the latter often is complete within microseconds to seconds, whereas the ribosome builds a protein at a rate of 1–20 residues per second, raising the potential for a more measured folding process that can lead to increased yield of native protein. The evolutionary significance of co-translational folding is supported by the intriguing observation that gene sequences tend to encode translational pauses – *via* codon choice, for example – a little downstream of domain boundaries. The suggestion, then, is that these pauses provide time for domains to fold without interference from subsequent portions of the chain. The higher rates of translation in prokaryotes and the inability to pause for long enough may contribute to the well-known difficulty of expressing active eukaryotic proteins in bacteria.

Gershenson and Gierasch take aim squarely at the limits of risks of reductionism, advocating a holistic approach to protein folding and function in the complex and compartmentalized cellular milieu. This vision is coming increasingly within reach, thanks to methodologic advances in, for example, microscopy and NMR. And the cellular environment is indeed complex. Chaperones facilitate folding by mechanisms that are still not fully understood, and specific cellular environments can be critical for folding, as proteins have evolved to folding in compartments with different redox states, different pH and different sets of chaperones. Macromolecular crowding may tend to drive folding and binding through simple sterics, but its effects become complex when one begins to account for weak attractive interactions among macromolecules and the kinetic consequences of viscosity. The cell's machinery for stabilizing marginally stable proteins may have profound evolutionary consequences, as this may be how novel or mutated proteins with adaptive functions but poor stability persist long enough to become fixed in the genome.

Bowie then takes on forces that drive folding, focusing on the role of hydrogen bonds in membrane protein folding and stability. Although hydrogen bonds *in vacuo* can be worth -5 kcal/mol or more, experiment shows that these bonds typically contribute much less,

perhaps -1 kcal/mol, to the stability of water-soluble proteins. Their relatively small contribution in this context presumably results in large part from the fact that intraprotein hydrogen bonds are in competition with water-protein hydrogen bonds. One might thus expect that hydrogen bonds would carry enormous weight in driving the assembly of membrane proteins within the non-polar membrane environment. However, rigorous double-mutant cycle studies indicate that hydrogen bonds are also weak in membrane proteins, surprisingly so. Bowie proposes that this observation is also explained by competition, but this time from alternative hydrogen-bonding partners within the protein, instead of from ambient water. And, even though individual hydrogen-bonds tend to be weak, their net effect can be large. Indeed, even non-traditional H-bonds in which alpha-carbons serve as weak donors appear to play a significant role as links between transmembrane helices.

Schreiber and Keating take us from folding to binding with a study of patterns of specificity and multi-specificity in protein-protein binding. The plasticity and versatility of proteins are highlighted by the ability of a protein family, or even a single protein, to interact with multiple partners in a highly controlled manner. Some proteins manage to bind multiple other proteins at essentially a single surface site, other proteins present multiple binding surfaces, while yet another set of proteins exploit an unstructured segment that adopts different conformations in order to interact with different proteins. Further diversity in protein-protein interactions derives from post-translational modifications, such as phosphorylation. In all cases, the authors emphasize that the biological significance of a binding affinity, whether weak or strong, can be assessed only in the context of the biologically relevant concentration of the proteins involved.

In their review article, Tzeng and Kalodimos delve further into the physics of binding and allostery, providing intriguing insights into microscopical mechanisms behind these phenomena gained by advanced NMR methods. First, they review cases where binding of a ligand at one site of protein elevates the statistical weight of conformations at a remote site that are competent to bind another ligand. These changes appear to be functionally relevant, even when the remote site conformations are still weakly populated (e.g. <10%). Advances in NMR methods can now provide structural details regarding such functionally important, though weakly occupied, states. Second, they discuss situations in which occupation of one binding site produces little change in the mean conformation at a second site, but instead produces significant changes in conformational fluctuations that mediate cooperativity between the two sites by altering the entropic consequences of binding.

Finally, Wang *et al.* offer a detailed account of the M2 proton channel from influenza virus, a compact, elegant model system in which these and other physical principles play out. This pH-activated, membrane-bound protein is targeted by the 'flu medications amantadine and rimantadine. The full channel is a homotetramer of a 97-residue protein, but much of the action focuses in a 19-residue transmembrane helical domain that tetramerizes, conducts protons, and binds the drugs. Substantial structural information is now available from X-ray and NMR experiments, which have revealed a pleasingly symmetrical "histidine box" centered on the channel axis and directly involved in the functional binding and release of protons. The structural data also indicate that the cationic drugs bind neatly in the channel lumen. The increasingly comprehensive picture that is emerging for the linked binding of drugs and protons by this protein bear directly upon efforts to overcome mutational resistance to drugs and serve as one example of how fundamental studies of folding and binding can help to prepare us for the challenge of pandemic influenza.

Taken together, the reviews presented in this issue convey how evolution has uncovered an astonishing range of versatility in protein mechanisms, taking advantage of the subtlest

physical phenomena – e.g. pauses in protein translation, weak hydrogen bonds, and long-ranged entropic effects – to generate both structure and function. This exploitation of every aspect and feature of a component is, perhaps, one of the hallmarks of evolved, as opposed to engineered systems. If so, we should continue to expect complexity and the unexpected as we probe how molecules form and interact in the processes of life.

Biographies

Mike is Professor and Chair in Computer-Aided Drug Design at the Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego. He received his bachelor's degree in Bioengineering from Harvard College in 1981 and his Ph.D. and M.D. from Columbia University. Mike's research centers on development of theory and computational methods, with recent interests focusing on configurational entropy and mechanical stress at the molecular level. His group also applies computational methods to applications in molecular design, from host-guest systems to small-molecule drugs and proteins. And the Gilson lab operates the BindingDB molecular recognition database, which provides access to experimental data on protein-small molecule and host-guest binding (www.bindingdb.org).

Sheena is currently Professor of Structural Molecular Biology and Deputy Director of the Astbury Centre for Structural Molecular Biology. She received her bachelor's degree in biochemistry from the University and her PhD from the University of Cambridge, UK. Sheena's research focuses on the determination of the molecular mechanisms of protein folding and misfolding using an array of experimental methods from biochemistry to biophysics, including NMR, mass spectrometry, kinetic analysis and single molecule methods. Systems currently under study include the bacterial immunity proteins, which have to fold rapidly and bind to their cognate colicins tightly to protect the host organism from intoxication, and the protein β_2 -microglobulin, a protein for which misfolding imparts the onset of aggregation and is the causative agent of the human disorder dialysis-related amyloidosis.