Published in final edited form as:

J Biomol Struct Dyn. 2013 ; 31(9): 1024–1027. doi:10.1080/07391102.2012.748547.

The unsolved “solved-problem” of protein folding

B. Montgomery Pettitt*

Sealy Center for Structural Biology, University of Texas Medical Branch, Galveston, TX
77555-0304, USA

The problem of protein folding presents a puzzle of such intrigue and depth that the scientific community cannot even agree on a single statement of the problem ((Ben-Naim, 2012) and references therein). The interplay between notions like hydrophobicity as the driving force or folding and what sort of free energy concepts and coordinates are required shows more intrigue than one normally finds in “solved problems”. A central figure whose hypothesis has defined a great body of interpretations and misinterpretations of the problem is that of Anfinsen (1973). The first issue that Prof. Ben-Naim takes on in his recent article (Ben-Naim, 2012) entitled “Levinthal’s question revisited, and answered” is the idea that Anfinsen’s thermodynamic hypothesis of protein folding has been misinterpreted, often, leading in part to interpretations of Levinthal’s paradox (Levinthal, 1968).

Some of the apparent confusion in the field with issues raised in even defining the protein folding problem is based on the words we employ and the concomitant concepts that are often used without sufficient precision. Prof. Ben-Naim goes to some lengths to define some of the relevant words and concepts needed from statistical thermodynamics. Central to this is the use of energy versus. free energy in describing the folding coordinate. It is true that many use these separate concepts interchangeably, which is incorrect. That free energy is the only relevant general, chemical driving force, rather than energy, which is not in question. When authors use terms like “energy landscape” rather than “free energy landscape” when describing a spontaneous process, it is incorrect no matter how common. Even Born occasionally slipped in this regard (Pettitt, 2000).

The extent of the arguments made by Prof. Ben-Naim are more subtle than just such simple issues as confusion between energy (or heat) and free energy. Besides the fact that free energy differences between states determines equilibrium concentrations or probabilities and so a system is incorrectly described as falling into one, even global minimum, he makes the point that adding a geometric dependence, such as a distance or folding coordinate, changes the mathematical description to that of a functional rather than just a function. This is a subtle but valid point. This is used to criticize the often poorly defined folding funnel concept. Others have also recently noted inconsistent use in the folding funnel literature as well (Karplus, 2011).

A point not often well appreciated is the physicochemical (structural or compositional) nature of the thermodynamic states in a folding process that one might want to describe with

*mpettitt@utmb.edu.

a given reaction coordinate, funnel or not. As an extreme, one could imagine the set of completely extended polymer geometries. Such states for proteins in water, which is a marginally poor protein solvent, would be highly unlikely (improbable), but there are an enormous number of relatively extended conformations, as often qualitatively described as the top of a folding funnel. If, however, one considers either chemically or modestly thermally denatured states, the number of probable configurations is far less and a radical funnel description would indeed be less appropriate. The competition of the polymer entropy with enthalpic and entropic contributions from solvent is appreciated in the primary literature (Bryngelson, Onuchic, Socci, & Wolynes, 1995) but not always clearly revealed in subsequent works as criticized here and elsewhere (Karplus, 2011). What the funnel might depict quantitatively is too often left to the imagination.

The second issue tackled in the Ben-Naim manuscript is the hydrophobic effect as the driving (free energy) mechanism in folding. The basis of his argument on the dominant, relevant protein system forces is well appreciated in that the interactions and consequent forces ensuring that the acidic and basic groups stay exposed to aqueous solvent are much stronger than those associated with the less soluble groups. The solvation of ionic groups indeed provides a strong thermodynamic driving force for their aqueous exposure. But is that enough to explain protein stability?

The text book dogma of folding is that hydrophobicity is the driving interaction of protein folding (Garrett & Grisham, 2002; Nelson & Cox, 2008; Voet, Voet, & Pratt, 2008). Much of the argument that protein folding is driven by hydrophobicity of the side chains goes back to the early measurements of Cohn and Edsall (1943); much was further formulated in the seminal review by Kauzmann (1959). The idea is to measure the free energies of transfer for the components of proteins (e.g. side chain surrogates, backbone mimics) between water and other solvents (Tanford, 1962). Like dissolves like, and the less polar side chains prefer a less polar environment, be it solvent or other like-polarity side chains. The protein backbone with its hydrogen bonding groups exposed has a favorable free energy of solvation (negative) but in this view “protected” from solvent by virtue of secondary and tertiary structure in folded proteins.

Of particular interest in this context is the transfer free energy difference for these model compounds of side chains and backbone between water and aqueous solutions containing either folding or unfolding cosolvents (Auton & Bolen, 2005; Tanford, 1964). To make sense of the thermodynamic data obtained for the models in relationship to globular proteins, it is helpful to include the effects of the change in exposed surface area due to conformational changes between the native and denatured state(s) (Auton & Bolen, 2005). This reveals the backbone solubility difference as being a central player rather than a passive, protected spectator (Rose, Fleming, Banavar, & Maritan, 2006; Street, Bolen, & Rose, 2006). While the backbone represents around a fourth of the total change in surface area, it makes a substantial, in many cases the dominant, contribution to the overall free energy change. This view is strengthened when the data are corrected for the solute activities (Auton, Holthauzen, & Bolen, 2007). Then the signal from the side chains is substantially diminished and that of the backbone remains dominant, although the details of the proportions depend on concentration (Auton, Rosgen, Sinev, Holthauzen, & Bolen, 2011).

A view that places the side chains and backbone solvation free energies as both having importance does not fit well with the classic dominant view of the hydrophobic effect on protein folding. The article being commented on here also takes a non-classical view of folding driving forces. As mentioned, it considers the importance of the solubilities of the charged side chains versus the less polar ones. While the ionic ones certainly make a contribution, the details of the decompositions of the experimentally determined transfer free energies (Auton & Bolen, 2007; Auton et al., 2007) do not support them as being more significant than the uncharged side chains in explaining protein stability.

Problems in defining hydrophobic interactions and hydrophobicity scales abound in the literature (Cornette et al., 1987). There is no consensus scale because there is no unique definition within the field of biochemistry. Surface science has a rigorous definition of the concept of the wetting–dewetting transition at the hard surface–liquid interface given to us by Young over two hundred years ago (Young, 1805). By this wetting contact angle definition many “hydrophobic side chains” of proteins are not hydrophobic in that they would not dewet.

Using the geometry of the surface or cavity in question adds an important dimension to the problem. Hydration is greatly affected by the surface concavity (Cheng & Rossky, 1998). Many would have difficulty responding to the recent counter intuitive experimental results of S. Martin on the sign of the hydration heat capacity in that rather peculiar series of hydrophobic binding events (Myslinski, DeLorbe, Clements, & Martin, 2011). The work and its interpretation are understood in terms of theoretical/computational analysis (Setny, Baron, & McCammon, 2010). The idea is that for a sufficiently sized and shaped a polar cavity, the water inside may be so disordered already that it in fact gains order when released to bulk. This takes the venerable Frank and Evans’ (Frank, 1945, Frank & Evans, 1945, Frank & Wen, 1957) idea of structure making and breaking into a new and unexpected thermodynamic realm. The phenomenon is interesting but far from universal in ligand binding. It is a sort of real but extreme limit.

Coming from another point of view, computer simulations (Hu, Lynch, Kokubo, & Pettitt, 2010; Tran, Mao, & Pappu, 2008) and experiments (Teufel, Johnson, Lum, & Neuweiler, 2011) on oligo-glycine in solution have shown that a system with no single equilibrium folded state reveals much about the role of the backbone. Oligomers of glycine 15–25 in length show collapse to a wet molten globule-like state in water, expand in urea, and contract further in TMAO. The collapse in water is related to the picomolar and worse solubility. The solubility gets progressively less as we proceed from Gly to Gly₂ to Gly_{few}, even though the dilute solvation free energy becomes increasingly favorable. The aqueous solutions undergo a phase separation turning cloudy with increasing concentration (Teufel et al., 2011). Thus, the initial collapse from an extended chain to a wet globule does not even require side chains; it requires only local concentration or chain length to be sufficient.

Many would agree that the two main issues in computational protein folding, as opposed to theories of folding, are the accuracy of the force fields and the issue of sampling the degrees of freedom. The community has made credible recent progress on both force fields (Ponder et al., 2010; Ren & Ponder, 2003) and sampling (Zhang & Ma, 2010; Zhang & Ma, 2012a).

The problem, even for practical engineering applications, still seems far less solved than a work in progress (Zhang & Ma, 2012b).

Both the philosophy and methods of science show differences in viewing problems from a target-based (phenomenological or top down) versus a caused-based (first principles or bottom up) approach. Based on the hydrophilicity of certain side chains Ben-Naim concludes that “there is overwhelming evidence that the hydrophobic effect, while important, is not a solution to Leventhal’s problem or the protein folding problem.” On this point, I am in agreement with Prof. Ben-Naim but not for his precise reasons. Phrasing the question of whether stable structures exist, “not because the proteins have some inner code that commands them to reach the target, but because the protein with this particular target 3D structure was selected by evolution” to my way of thinking simply clouds the issue of the role of molecular biophysics in biology. It seems that funnels and hydrophobicity need either precision or updating.

In the hydrophobicity case, new concepts have been put forth (Hu et al., 2010). In particular, the mixing of approximate electrostatics and an arbitrary hydrophobicity model has been replaced with terms from the underlying model Hamiltonian in the calculations: electrostatics plus van der Waals. Such a simple two-part categorization has the advantage of being well defined (Kokubo, Hu, & Pettitt, 2011). The fact that the van der Waals term includes both excluded volume (cavity formation) as well as the dispersion attractions makes this a flexible alternative to accessible surface area-based hydrophobicity “corrections”.

In this analysis, it seems the language of solution properties, and most importantly, solubilities provide a reliable (even historical) means for discussing the subject of protein folding. The funnel has been compelling but like hydrophobicity, hard to uniquely quantify and easy to misuse. The use of free energy changes and quantitative model decompositions provides a sound basis for linking the biology, chemistry, and physics of protein folding.

Acknowledgement

The author thanks Dr. Gillian Lynch for critically reading the manuscript. Support is appreciated from the Robert A. Welch Foundation and from the NIH via GM037657.

References

- Anfinsen CB. Principles that govern the folding of protein chains. *Science*. 1973; 181:223–230. [PubMed: 4124164]
- Auton M, Bolen DW. Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102:15065–15068. [PubMed: 16214887]
- Auton M, Bolen DW. Application of the transfer model to understand how naturally occurring osmolytes affect protein stability. *Methods in Enzymology*. 2007; 428:397–418. [PubMed: 17875431]
- Auton M, Holthauzen LMF, Bolen DW. Anatomy of energetic changes accompanying urea-induced protein denaturation. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104:15317–15322. [PubMed: 17878304]

- Auton M, Rosgen J, Sinev M, Holthauzen LM, Bolen DW. Osmolyte effects on protein stability and solubility: A balancing act between backbone and side-chains. *Biophysical Chemistry*. 2011; 159:90–99. [PubMed: 21683504]
- Ben-Naim A. Levinthal's question revisited, and answered. *Journal of Biomolecular Structure & Dynamics*. 2012; 30:113–124. [PubMed: 22571437]
- Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins: Structure, Function, and Bioinformatics*. 1995; 21:167–195.
- Cheng YK, Rossky PJ. Surface topography dependence of biomolecular hydrophobic hydration. *Nature*. 1998; 392:696–699. [PubMed: 9565030]
- Cohn E, Edsall J. Proteins, amino acids and peptides as ions and dipolar ions. American Chemistry Society Monograph Series. 1998; 392:696–699.
- Cornette JL, Cease KB, Margalit H, Spouge JL, Berzofsky JA, DeLisi C. Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. *Journal of Molecular Biology*. 1987; 195:659–685. [PubMed: 3656427]
- Frank HS. Free volume and entropy in condensed systems i. general principles. fluctuation entropy and free volume in some monatomic crystals. *Journal of Chemical Physics*. 1945; 13:478–492.
- Frank HS, Evans MW. Free volume and entropy in condensed systems. III. Entropy in binary liquid mixtures; partial molal entropy in dilute solutions; structure and thermodynamics in aqueous electrolytes. *Journal of Chemical Physics*. 1945; 13:507–532.
- Frank HS, Wen WY. Ion-solvent interaction. structural aspects of ion-solvent interaction in aqueous solutions: A suggested picture of water structure. *Discussions of the Faraday Society*. 1957; 24:133–140.
- Garrett, RH.; Grisham, CM. Principles of biochemistry: With a human focus. Belmont, CA: Harcourt College; 2002.
- Hu CY, Lynch GC, Kokubo H, Pettitt BM. Trimethylamine N-oxide influence on the backbone of proteins: An oligoglycine model. *Proteins: Structure, Function, and Bioinformatics*. 2010; 78:695–704.
- Karplus M. Behind the folding funnel diagram. *Nature Chemical Biology*. 2011; 7:401–404.
- Kauzmann, W. Some factors in the interpretation of protein denaturation. In: Anfinsen, CB.; Anson, ML.; Bailey, Kenneth; Edsall, John T., editors. *Advances in Protein Chem*. Vol. 14. Academic Press Inc; 1959. p. 1-63.
- Kokubo H, Hu CY, Pettitt BM. Peptide conformational preferences in osmolyte solutions: transfer free energies of decaalanine. *Journal of the American Chemical Society*. 2011; 133:1849–1858. [PubMed: 21250690]
- Levinthal C. Are there pathways for protein folding. *Journal de Chimie Physique*. 1968; 65:44.
- Myslinski JM, DeLorbe JE, Clements JH, Martin SF. Protein-ligand interactions: thermodynamic effects associated with increasing nonpolar surface area. *Journal of the American Chemical Society*. 2011; 133:18518–18521. [PubMed: 22007755]
- Nelson, DL.; Cox, MM. *Lehninger principles of biochemistry*. New York, NY: W.H. Freeman; 2008.
- Pettitt BM. A Perspective on “volume and heat of hydration of ions”. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*. 2000; 103:171–172.
- Ponder JW, Wu C, Ren P, Pande VS, Chodera JD, Schnieders MJ, DiStasio RA Jr. Current status of the AMOEBA polarizable force field. *The Journal of Physical Chemistry*. 2010; B114:2549–2564. [PubMed: 20136072]
- Ren P, Ponder JW. Polarizable atomic multipole water model for molecular mechanics simulation. *The Journal of Physical Chemistry*. 2003; B107:5933–5947.
- Rose G, Fleming P, Banavar J, Maritan A. A backbone-based theory of protein folding. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103:16623–16633. [PubMed: 17075053]
- Setny P, Baron R, McCammon JA. How can hydrophobic association be enthalpy driven? *Journal of Chemical Theory and Computation*. 2010; 6:2866–2871. [PubMed: 20844599]

- Street TO, Bolen DW, Rose GD. A molecular mechanism for osmolyte-induced protein stability. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; A103:13997–14002. [PubMed: 16968772]
- Tanford C. Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. *Journal of the American Chemical Society*. 1962; 84:4240–4247.
- Tanford C. Isothermal unfolding of globular proteins in aqueous urea solutions. *Journal of American Chemical Society*. 1964; 86:2050–2059.
- Teufel DP, Johnson CM, Lum JK, Neuweiler H. Backbone-driven collapse in unfolded protein chains. *Journal of Molecular Biology*. 2011; 409:250–262. [PubMed: 21497607]
- Tran HT, Mao A, Pappu RV. Role of backbone-solvent interactions in determining conformational equilibria of intrinsically disordered proteins. *Journal of the American Chemical Society*. 2008; 130:7380–7392. [PubMed: 18481860]
- Voet, D.; Voet, JG.; Pratt, CW. *Principles of biochemistry*. New York, NY: Wiley; 2008.
- Young T. An essay on the cohesion of fluids. *Philosophical Transactions of the Royal Society*. (London). 1805; 95:65.
- Zhang C, Ma J. Enhanced sampling and applications in protein folding in explicit solvent. *The Journal of Physical Chemistry*. 2010; 132:244101-1–244101-7.
- Zhang C, Ma J. Estimating statistical distributions using an integral identity. *The Journal of Physical Chemistry*. 2012a; 136:204113-1–204113-9.
- Zhang C, Ma J. Folding helical proteins in explicit solvent using dihedral-biased tempering. *Proceedings of the National Academy of Sciences of the United States of America*. 2012b; A109:8139–8144. [PubMed: 22573819]