Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 2426–2427, March 1995

Commentary

Biomolecular folding *in vacuo*!!!(?)

Peter G. Wolynes

School of Chemical Sciences, University of Illinois, Urbana, IL 61801

At the molecular level, life, as we know it, goes on in water. A liquid milieu is doubtless essential to provide rapid mobility to both the small molecules whose traffic is so essential to life and to the functioning macromolecules which act upon them. The omnipresence of water has exerted a profound influence on the evolution of biomolecular function, specificity, and stability. The self-assembly, or folding, of protein molecules in particular has been thought to involve water most intimately. Therefore, the observations reported by McLafferty and coworkers (1) of processes resembling the folding and unfolding of cytochrome *c* ions in vacuum immediately commands our attention and a careful examination of our views on the dominant forces involved in protein folding.

While theorists have studied proteins *in vacuo* for some time (2), experimental protein chemistry of any kind in the vacuum may seem novel. At equilibrium, proteins have low vapor pressures; they are hardly "soluble" in the vacuum. Nevertheless a variety of techniques such as electrospray and laser deposition can inject a few nanomoles of protein into the gas phase. Thus, mass spectrometry is now routinely used to sequence biomolecules (3). The high sensitivity and specificity of mass spectrometry have also made it useful for characterizing the state of proteins taken from solution. Beautiful work by Dobson and collaborators (4) uses mass spectrometry to characterize various folding intermediates in solution by exploiting solution-phase proton exchange to label which residues are exposed to solvent. Studying the dynamics of what happens to a biomolecule once in the vacuum is a difficult matter, however, since there are so few molecules available. In the study of McLafferty and co-workers (1), chemical reactions and lasers come to the rescue; they characterize the conformations of the protein *in vacuo* using gas-phase hydrogen–deuterium exchange, a process stoichiometrically analogous to the very different chemistry used in solution. Again the number of exchangeable protons should give an idea of the openness of the protein conformation during gas-phase folding. Since the molecules are held for a long time in a Fourier transform mass spectrometer, McLafferty and co-workers also can use infrared lasers to heat the molecules. This apparently causes them to unfold *in vacuo*. Collisions with other gas molecules can change the charge on the protein ions, thereby allowing conformational changes. This is analogous to the way pH changes modulate folding in solution. In each case the high sensitivity of mass spectrometry allows the monitoring of very subtle changes in a very small number of molecules so that kinetic studies can be carried out. These studies indicate that gas-phase folding involves not just two states but several different populations of protein molecules that display varying degrees of proton exchangeability and, therefore, different degrees of openness. McLafferty and co-workers convincingly show both reversible and irreversible thermal transitions between these different conformational states. Whether these transitions involve the same states as the folding in solution is a question still worthy of discussion. The thermodynamic and kinetic stability and uniqueness of the collapsed structures all need to be considered.

Historically various "forces" have been thought to be important in protein folding. Textbooks highlight three effects in discussing protein stability (5). These are hydrogen bonding (6), hydrophobic forces (7), and packing (8). Each of these plays a role in the architecture of fully folded proteins. The local secondary structures, α-helix and the β-sheet, are quite obvious features of proteins and are dominated by hydrogen bonds. Nearly equally prominent in protein architecture is the segregation of apolar hydrophobic and polar hydrophilic residues into the interior and surface of the protein, respectively. It is also quite apparent that large cavities are exceptional in proteins. Proteins are nearly as close packed as amino acid crystals, so the repulsive forces and steric constraints are very important. Both hydrogen bonding and the segregation of polar and apolar residues should be affected strongly by the presence or absence of water, since the water can provide alternative hydrogen bonds and solvate the charges on the polar residues.

From the point of view of hydrogen bonding, the stability of the native protein *in vacuo* is not compromised. Without the alternative of hydrogen bonding to the solvent, the classical secondary structures should be even more thermodynamically stable. Indeed, their presence was predicted theoretically by Pauling without any concern for solvation phenomena. On the other hand, at first glance the contribution of hydrophobic forces to stability would seem to be a problem *in vacuo*. The vacuum being very apolar can be thought of as a very hydrophobic medium much like the lipid of a biomembrane. From this viewpoint it would seem sensible for a protein to turn inside out with its hydrophilic or polar residues inside and hydrophobic apolar residues outside. A transition to such a kind of collapsed structure cannot be ruled out so far by the experiment reported in this issue (1). It is worth bearing in mind, however, that the hydrophobic force is actually a composite of many microscopic forces. Microscopically one contribution to stability is the ordinary van der Waals attraction between the residues of the polypeptide. This is still present *in vacuo*. A separate contribution comes from van der Waals attractions between the water molecules and the protein atoms and between the waters themselves. Privalov has argued on the basis of microcalorimetry studies of solution-phase folding transitions that most of the stabilization, in fact, comes from the intramolecular van der Waals forces themselves and that the residual interactions with water are actually destabilizing (9). On this basis we would expect the folded protein structure to be at least metastable. This would be in accord with the many computer simulations of proteins *in vacuo* that usually hold together for a few nanoseconds of real time. Could the protein in its native solution structure be *globally* thermodynamically stable *in vacuo*; that is, are there no lower-energy conformations? The simple hydrophobic vacuum picture would argue otherwise! It is possible, however, that packing forces could provide all of the needed specificity. Richards (8), Chan and Dill (10), and many others to varying degrees have argued that many (but perhaps not all) structural features of folded proteins can be described in what statistical mechanicians would call a "van der Waals" picture. In this picture a relatively nonspecific attractive force causes the protein molecule to become compact. When the density is high enough, excluded volume and packing forces can then cause a specific ordered structure to arise, much like the crystallization of hard spheres

(11). Similarly, the knobby shapes of the amino acid side chains might be able to fit together efficiently in only one way. However, most computational chemists would probably be rather surprised if the minimum-energy vacuum and solution structures agreed. The local flexibility of protein molecules as shown with most molecular mechanics energy functions allows van der Waals forces and charge interactions to be quite well satisfied, even by conformations with far from native topologies (12, 13). Molecular mechanics gives a rough energy landscape for the protein molecule *in vacuo*. Here many other structures would be competitive with the x-ray structure energetically. These calculations suggest that the differential solvation effects of water are essential in selecting the unique native structure of proteins. The experiments themselves of McLafferty and co-workers suggest that there may be a significant difference between the least exchangeable, most folded or compact species in the gas phase and the native structure in solution, since their number of protons available for exchange differ.

Could the gas-phase folded species differ from the solution native structure but still be a unique conformation? If, indeed, gas-phase energies differ very much from the solution free energies, this would also be a significant surprise. The native protein sequence, which has evolved to have a nearly unique, kinetically accessible, low-free-energy conformation in solution, would really be a random sequence as far as folding *in vacuo*. The property of having a kinetically accessible unique ground state is not common to all sequences but requires the interactions encoded in the sequence of the protein to be in harmony with each other in the global minimum structure (ref. 14 and references therein). This is known as the principle of minimal frustration. If the van der Waals packing picture were wrong and if differential solvation forces were important in giving selectivity, it seems unlikely that the sequence of cytochrome *c* would satisfy this principle for folding in the vacuum since cytochrome evolved in water. Thus, the lowest state would probably not be kinetically accessible. The large multiplicity of different conformations found by McLafferty and co-workers would not be in disagreement with this.

Definitively establishing the degree of structural similarity between the compact structures of McLafferty and co-workers and the native solution structure thus becomes a very high priority, both to understanding the microscopic forces of folding and perhaps to understanding evolutionary constraints on protein-folding kinetics. McLafferty and co-workers note that tandem mass spectroscopy might establish which protons are exchanged *in vacuo* and thereby tell us the similarity of the structures. Convincing evidence that the compact structures reported by the Cornell chemists correspond with the native solution structure would be impressive confirmation of the van der Waals/steric picture as the dominant paradigm for the specificity of protein folding. If *in vacuo* folding differs from folding in solution as expected when using molecular-mechanics energy functions, it will be interesting to study how much water is needed to allow normal folding to go on. Recent computer simulations by Steinbach and Brooks (15) suggest that the addition of only a few tens of water molecules is necessary to give the dominant solvation effects. Experiments by Doster *et al.* (16) establish as well that protein function can be restored in desiccated protein powders by the addition of considerably less than a single monolayer of water. Studying folding *in vacuo* by the stepwise addition of waters will certainly be quite difficult, but potentially very rewarding.

The experiments of McLafferty and coworkers (1) raise the prospect of allowing us to understand the physics of the self assembly of chain molecules in a very exotic environment. Even if natural proteins, having evolved in an aqueous environment, do not fold to unique structures *in vacuo*, it is not at all obvious that molecules that do so cannot exist. Speculations about astrobiochemistry and biomolecules in interstellar clouds have long been with us (17). If such molecules, in fact, never have evolved, the challenge of deliberately designing a molecule to fold in vacuum is still far from unthinkable. Great progress in designing proteins that fold in the hydrophobic environment of membranes has already been achieved. With such designed macromolecules, the powerful high-resolution spectroscopic methods possible in the gas phase could be used to illuminate the problem of the self-assembly.

1. Wood, T. D., Chorush, R. A., Wampler, F. M., III, Little, D. P., O'Connor, P. B. & McLafferty, F. W. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 2451–2454.
2. McCammon, J. A. & Harvey, S. C. (1987) *Dynamics of Proteins and Nucleic Acids* (Cambridge Univ. Press, Cambridge, U.K.).
3. Biemann, K. & Scobel, H. A. (1987) *Science* **237,** 992–998.
4. Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T. & Dobson, C. M. (1993) *Science* **262,** 896–900.
5. Branden, C. & Tooze, J. (1991) *Introduction to Protein Structure* (Garland, New York).
6. Pauling, L., Corey, R. B. & Branson, H. R. (1951) *Proc. Natl. Acad. Sci. USA* **37,** 205–211.
7. Kauzmann, W. (1959) *Adv. Protein Chem.* **14,** 1–57.
8. Richards, F. M. (1974) *J. Mol. Biol.* **82,** 1–14.
9. Privalov, P. L. & Gill, S. J. (1988) *Adv. Protein Chem.* **39,** 193.
10. Chan, H. S. & Dill, K. A. (1993) *Phys. Today* **46,** 24–32.
11. Chandler, D., Weeks, J. D. & Anderson, H. C. (1983) *Science* **220,** 787–794.
12. Novotny, J., Brucade, R. & Karplus, M. (1984) *J. Mol. Biol.* **177,** 787.
13. Wang, Y., Zhang, H., Li, W. & Scott, R. A. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 709–713.
14. Bryngelson, J., Onudric, J. N., Socci, N. & Wolynes, P. G. (1995) *Proteins Struct. Funct. Genet.*, in press.
15. Steinbach, P. J. & Brooks, B. R. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 9135–9139.
16. Doster, W., Bachleitner, A., Dunau, R., Hiebl, M. & Lüscher, E. (1986) *Biophys. J.* **50,** 213.
17. Hoyle, F. (1957) *The Black Cloud* (Harper, New York).