Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 223 (2003) 263–265
Journal of Theoretical Biology
www.elsevier.com/locate/jtbi

Letter to the Editor

A comment on the protein folds as platonic forms

Denton et al. (2002) have presented compelling evidence that protein folds ought to be understood as arising from physical laws rather than natural selection. Furthermore, they suggest this could have "implications regarding the origin of proteins, the origin of life and the fundamental nature of organic form." They do not, however, explain what the physical basis is for understanding the origin of protein folds. Here, we wish to address this key missing ingredient.

Proteins are linear chain molecules made of amino acids. Under physiological conditions, they fold reproducibly and rapidly into somewhat compact forms called their native state structures. For proteins, form determines function. The total number of distinct protein folds are only a few thousands in number, these folds have as their building blocks elegant helices and almost planar sheets and protein structures are flexible and versatile (Denton et al., 2002). These features cannot be explained using conventional ideas from polymer physics. Compact polymers do not typically exhibit secondary motifs and the total number of compact structures is astronomically large, features that are in accord with those found in simple model systems comprised of spheres tethered along a chain.

We have recently shown (Banavar and Maritan, 2003; Banavar et al., 2003a) that the proper way of describing a chain molecule is by incorporating the inherent anisotropy—each particle of the chain has associated with it a local direction defined by its neighboring particles along the chain. This effect is captured by viewing the chain molecule, not as a balls and string model, but as a chain of disks or equivalently as a tube of non-zero thickness. The axis of the tube is the chain and the non-zero thickness provides room for the atoms of the amino acids in order to avoid steric overlaps (Ramachandran and Sasisekharan, 1968). (Surprisingly, in the continuum limit, one needs to discard pairwise interactions and work with suitable three-body interactions in order to describe the self-avoidance of such a tube. As a bonus, this description eliminates singularities in the interaction potential and the absolute need for use of renormalization group techniques (Banavar et al., 2003b).)

Let us consider, for simplicity, a homopolymer chain molecule with no amino acid specific details added in—our goal is to understand the common features of all proteins. One may write a Hamiltonian of the form

$$H_{chain} = \sum_i u(r_{i,i+1}) + \sum_{i<j} V_2(r_{i,j}) + \sum_{i<j<k} V_3(r_{i,j,k}), \quad (1)$$

where $r_{i,j}$ denotes the distance between particles i and j, $r_{i,j,k}$ is the radius of a circle going through particles i, j and k. u is a generic tethering potential (for example, one which constrains $r_{i,i+1}$ to be constant), V_2 has an attractive component which is availed of when two particles are within a threshold distance of R_1 from each other (V_2 captures the role of hydrophobicity in promoting compaction) and V_3, the tube constraint, is a hard core potential which forbids any three body radius from being smaller than R_0, the tube thickness (Gonzalez and Maddocks, 1999).

In proteins, there is a rich interplay between the range of attractive interactions, R_1, and the tube thickness, R_0 (Banavar et al., 2002a). The two length scales are comparable because they are both determined by the side chains of the amino acids: there are Angstrom scale interactions between the outer atoms of the side chains as the water is squeezed out of the core of the protein and the tube thickness is controlled by the steric avoidance of these same side chains. Fig. 1 shows a schematic sketch of the number of contacts in the ground state of a short tube (or equivalently the negative of the ground state energy) when one varies the ratio of R_0 to R_1. For small tube thickness, one obtains a phase analogous to the generic compact polymer phase whereas, for large tube thickness, one gets a swollen phase within which the tube is too fat to avail of the attraction. There is a phase transition between these two phases when $R_0 \sim R_1$. In the vicinity of this transition, short tubes fold in a marginally compact way into helices of a specific pitch to radius ratio (Maritan et al., 2000) (quantitatively comparable to that observed in proteins) and almost planar sheets.

This novel phase of matter (Banavar and Maritan, 2003; Banavar et al., 2003a) has been exploited by Nature to house protein folds because of many of its

*Tel.: +1-814-8631089; fax: +1-814-865-0978
†Tel.: +39-040-2240462; fax: +39-040-3787528.

0022-5193/03/$ - see front matter
doi:10.1016/S0022-5193(03)00098-5

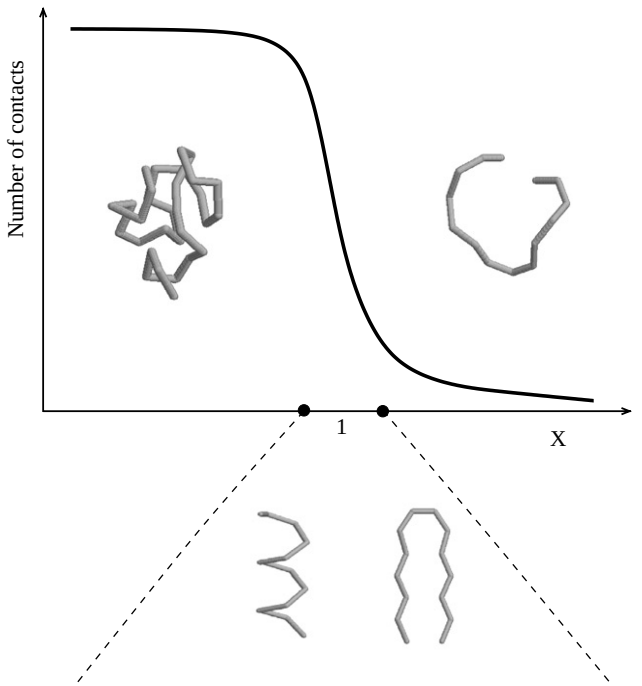

Fig. 1. Schematic sketch of the number of contacts in the ground state of a tube subject to compaction versus X, the ratio of the tube thickness to the range of the attractive interaction. For small X, one obtains a highly degenerate compact phase whereas one gets a swollen phase at large X. There is a phase transition between these two phases around $X \sim 1$. Typical ground state conformations of short tubes in each of the phases are shown. For more details, see Banavar and Maritan (2003).

advantages including those described by Denton et al. (2002) and Finkelstein and Ptitsyn (2002):

(1) The presence of a limited number of folds arises because of the constraints of the tube anisotropy. In the marginally compact phase, one requires that a space-filling conformation is created (in order to expel water from the hydrophobic core) with a preferentially parallel orientation of nearby tube segments.

(2) The native state folds may be thought of as "pre-existing Platonic molds" (Denton et al., 2002; Denton and Marshall, 1999) that a sequence must choose from for its native state structure.

(3) The limited number of folds that a sequence can adopt explains the relatively large basin of attraction (in a dynamical sense) for each of them. The thousand or so putative native structures allow for both diversity and stability, which are dual characteristics required for evolution to be successful (Anderson, 1983).

(4) The marginally compact phase exists in the vicinity of a phase transition which accounts for the exquisite sensitivity of protein conformations to the right types of perturbations.

(5) Because the set of protein folds are predetermined by physical law, the sequence and the functionality evolve in order to best make use of these folds.

(6) Proteins are able to fold dynamically in an all-or-none transition into the native state (Bogatyreva and Finkelstein, 2001). This can be accounted for in the tube picture by noting that in the marginally compact phase, the energy scale of interaction is relatively weak and therefore the transition temperature is low and entropic effects are not important. The careful orientational positioning required of nearby tube segments (Banavar et al., 2002b) leads to the tube snapping into its native state structure.

(7) The tube picture explains in a simple way the formation of fibril-like structures called amyloids which are implicated in a variety of diseases (Dobson, 2002).

The situation here is reminiscent of a much older problem of the prediction of crystal structures. Beautiful mathematical ideas pertaining to symmetry and geometry can be used to enumerate possible crystal structures. Given a material, one can then ask which structure it would fit best given the details of the interatomic interactions. In similar spirit, the protein folds are predetermined by considerations of physical law, and more specifically, the geometry of marginally compact tubes, and different amino acid sequences choose from among these structures to house their native state.

We are intrigued by the suggestion of Denton et al. (2002) that "the lawful nature of the folds together with the intriguing fact that many of the 20 protogenic amino acids—out of which the folds are constructed—are amongst the most common amino acids found in meteorites and the easiest amino acids to generate in pre-biotic syntheses is surely of considerable significance, consistent with and supporting a deterministic theory of the origin of life (or at least of proteins) and by extrapolation the whole Platonic cosmogony—raising the possibility that all organic forms and indeed the whole pattern of life may finally prove to be the determined end of physics and life a necessary feature of the fundamental order of nature."

We are indebted to our collaborators Alessandro Flammini, Davide Marenduzzo, Cristian Micheletti, Flavio Seno and Antonio Trovato for valuable discussions. This work was supported by Confinanziamento MURST, INFM, NASA and the Penn State MRSEC under NSF Grant DMR-0080019.

References

Anderson, P.W., 1983. Suggested model for prebiotic evolution: the use of chaos. Proc. Natl Acad. Sci. USA 80, 3386–3390.

Banavar, J.R., Maritan, A., 2003. Geometrical approach to protein folding: a tube picture. Rev. Mod. Phys. 75, 23–34.

Banavar, J.R., Maritan, A., Micheletti, C., Trovato, A., 2002a. Geometry and physics of proteins. Proteins 47, 315–322.

Banavar, J.R., Maritan, A., Seno, F., 2002b. Anisotropic effective interactions in a coarse-grained tube picture of proteins. Proteins 49, 246–254.

Banavar, J.R., Flammini, A., Marenduzzo, D., Maritan, A., Trovato, A., 2003a. Geometry of compact tubes and protein structures. ComPlexUs 1, 4–13.

Banavar, J.R., Gonzalez, O., Maddocks, J.H., Maritan, A., 2003b. Self-interactions of strands and sheets. J. Stat. Phys. 110, 35–50.

Bogatyreva, N.S., Finkelstein, A.V., 2001. Cunning simplicity of protein folding landscapes. Protein Eng. 14, 521–523.

Denton, M.J., Marshall, C.J., 1999. Laws of form revisited. Nature 410, 417.

Denton, M.J., Marshall, C.J., Legge, M., 2002. The protein folds as Platonic forms: new support for the pre-Darwinian conception of evolution by natural laws. J. theor. Biol. 219, 325–342.

Dobson, C.M., 2002. Protein-misfolding diseases: getting out of shape. Nature 418, 729–730.

Finkelstein, A.V., Ptitsyn, O.B., 2002. Protein Physics: A Course of Lectures. Academic Press, Boston.

Gonzalez, O., Maddocks, J.H., 1999. Global curvature, thickness and the ideal shapes of knots. Proc. Natl Acad. Sci. USA 96, 4769–4773.

Maritan, A., Micheletti, C., Trovato, A., Banavar, J.R., 2000. Optimal shapes of compact strings. Nature 406, 287–290.

Ramachandran, G.N., Sasisekharan, V., 1968. Conformations of polypeptides and proteins. Adv. Protein Chem. 23, 283–438.

Jayanth R. Banavar*
Department of Physics, 104 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA
E-mail address: jayanth@phys.psu.edu

Amos Maritan†
International School for Advanced Studies (S.I.S.S.A.), Via Beirut 2-4, 34014 Trieste, INFM and the Abdus Salam International Center for Theoretical Physics, Trieste, Italy
E-mail address: maritan@sissa.it