Hypothesis

Is tensegrity a unifying concept of protein folds?

Giuseppe Zanotti[a,b,*], Concettina Guerra[c]

[a] *Department of Organic Chemistry, University of Padua, Via Marzolo 1, 35131 Padua, Italy*
[b] *Venetian Institute of Molecular Medicine, Via Orus 2, Padua, Italy*
[c] *Department of Electronics and Informatics, University of Padua, Via Gradenigo 6a, 35131 Padua, Italy*

Received 6 June 2002; revised 1 November 2002; accepted 19 November 2002

First published online 19 December 2002

Edited by Thomas L. James

Abstract **We suggest that the three-dimensional architecture of globular proteins can be described in terms of tensegrets, i.e. structural elements that are held together through attractive and repulsive forces. Hard elements of tensegrets are represented by secondary structure elements, i.e. α-helices and β-strands, while the role of elastic elements is played by attractive and repulsive atomic forces. Characteristics of tensegrets is that they can auto-assemble and that they respond to changes of tension in some part of the entire object through a deformation in another part, thus partially preserving their structure, despite their deformation. This latter property well explains both the folding process and the behavior of globular proteins under mild denaturing conditions, as revealed by the molten globule state.**

Key words: Tensegrity; Geodesic; Protein folding; Protein structure; Molten globule state

1. Tensegrity and geodesic symmetry

Tensegrity is a neologism used to indicate the balance among tension and repulsion (compression) that causes some objects to maintain their structural integrity under tension (for a general review of the subject, see for example [1]). The word tensegrity, originally devised by Buckminster Fuller to designate some special kind of sculptures, rapidly spread into other fields and it is now applied to several different objects, from buildings to cells, tissues or supramolecular aggregates. A structure characterized by tensegrity can be thought of as made up of rigid elements (for example bars) held together by elastic members (for example elastic tendons) (Fig. 1). More generally, according to the original Fuller definition [2], the structural shape of a tensegret is guaranteed by the interaction between a set of members in tension and a set of members in compression: it is the balance between these opposite forces that makes the structure intrinsically stable. It must be noticed that elements that exert attraction (compression) and others that exert repulsion (tension) are not necessarily distinguished, since the same element can exert both attraction and repulsion, according to the situation. Important features of objects obeying the tensegrity principles are their ability to auto-assemble and their relative flexibility: an increase of tension in one part of the object causes a deformation that increases compression in another part. So they have the ability to partially preserve their structure despite being deformed.

Tensegrity can be observed at different levels, from the macroscopic world to the atomic scale. During the past years, work has been done to demonstrate that some mechanical behaviors of living cells can be explained in terms of tensegrity [3–7] and perhaps their origin can be described using the tensegrity model [8].

A concept related to tensegrity is that of geodesic structures. A large number of points on a spherical surface symmetrically distributed in order to minimize variations in the distance between neighboring points gives rise to a so-called geodesic object. For a perfect symmetry only 12 points can be accommodated on a spherical surface (icosahedral symmetry), but many more points can be located assuming a quasi-symmetry (geodesic subdivision). Ideal tensegrets generally can be described by some sort of geodesic symmetry. The classic example of a geodesic tensegret is perhaps represented by a real molecule, fullerene (Fig. 2a), an allotropic state of carbon [9], but other examples of more complex molecular organizations can be listed, for example the capsid of spherical viruses (Fig. 2b) [10]. More complex supramolecular organizations can be described in terms of tensegrets, like that displayed in Fig. 2c, where an idealized, theoretical model of the orga-

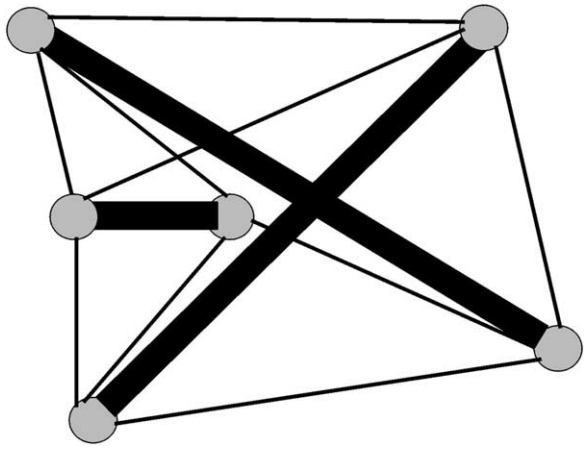

Fig. 1. An idealized representation of a superstable tensegret, i.e. a manufact built to obey the theoretical principles of tensegrity (adapted from [1]). Bold lines represent 'hard' elements, i.e. bars, thin lines 'soft' elements, i.e. elastic tendons.

*Corresponding author. Fax: (39)-49-8275239.
E-mail address: giuseppe.zanotti@unipd.it (G. Zanotti).

0014-5793/02/$22.00 © 2002 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/S0014-5793(02)03853-X

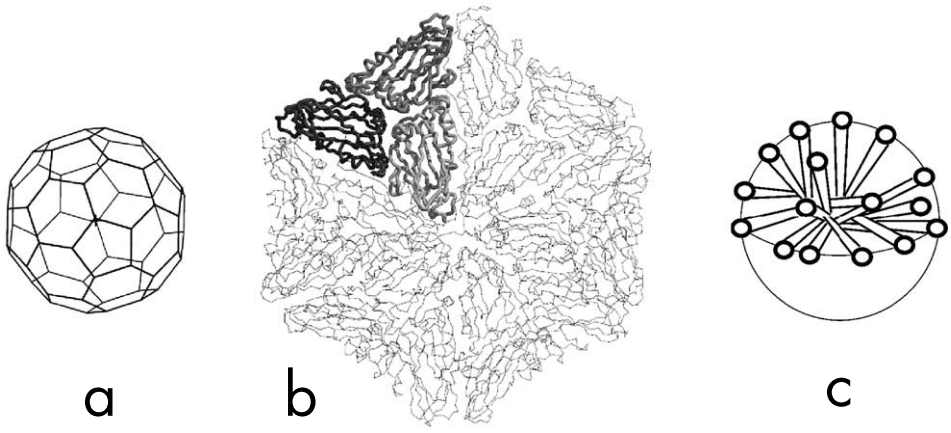

Fig. 2. a: C60 Fullerene, an allotropic state of carbon made up of 60 C atoms. Each carbon atom is bonded to three other carbon atoms in an infinite two-dimensional array. This gives rise to some sort of soccer ball structure, whose shape and symmetry recall the geodesic domes popularized by Buckminster Fuller. b: Cα chain trace representation of the capsid of southern bean mosaic virus (only the portion of the capsid towards the reader is shown) [25]. Twenty copies of three polypeptide chains with similar folds are arranged in a nearly spherical shell with icosahedral symmetry. c: Theoretical model of a possible organization of amphipathic molecules in a spherical micelle. Spheres represent hydrophilic heads, cylinders the hydrophobic tails (adapted from [11]). This kind of geodesic congeries preserves approximate symmetry.

nization of amphipathic molecules in a micelle is represented: hydrocarbon tails appear, at first sight, randomly distributed, but in fact they preserve a rough 'order' or 'symmetry', that can be described by the so-called geodesic congeries [11].

2. Tensegrity and globular proteins

Is there any relationships between tensegrity and globular protein organization? It is well known that the huge number of three-dimensional structures of globular proteins nowadays available (more than 16 000 structures deposited at the end of 2001 in the Protein Data Bank [12]) can be grouped in a quite limited number of basic folds (see, for example, [13]). The molecular models present in the database are not fully representative of all the protein classes encoded by a single genome: the structures determined via X-ray diffraction are strongly influenced by the possibility of being crystallized. Furthermore, membrane proteins are under-represented in the set, while nuclear magnetic resonance structures are on the contrary limited to small-size proteins. Despite that, there is a general consensus that the total number of basic folds is definitely limited. For example, they can be grouped into a few hundred 'unique' folds [14] and reasonable estimates indicate their total number to range from about 1000 to 5000 [15,16]. The hypothesis that protein folds found in nature represent a finite set of platonic forms has also been put forward [17]. Moreover, from the examination of the unique folds detected so far, local similarities or regularities have often been noticed: for example, it has been long recognized that α-helices tend to pack in such a way that their axes form preferred angles, giving rise to a limited number of orientations [18–21]. β-strands in a sheet are also naturally oriented in a preferred way.

The structure of a globular protein is classically described using a hierarchical scheme, i.e. it is considered as having a secondary, tertiary and possibly quaternary structure. Many data in the literature suggest that this hierarchical organization is not simply due to our abstraction, but that some secondary structural elements, in particular α-helices, play an active role in the folding process: in some cases it has been demonstrated that they represent the 'nucleation site' from which the tertiary structure evolves (for a review on folding, see [22]). Besides, it is well established that for most globular proteins the denaturation process goes through an intermediate state(s), called 'molten globule state' [23]. In a molten globule state the three-dimensional structure becomes distorted, but most or all of the secondary structure elements are preserved, despite being now organized in space in a relatively different way with respect to the 'native' structure.

We must remember that the conformation of a protein molecule is stabilized by a subtle equilibrium between attractive and repulsive forces. This equilibrium is in some way dynamic, i.e. van der Waals forces can exert an attraction or a repulsion, according to the position of atoms in space: an attractive force can become a repulsive one when two atoms come too close. This is also true for electrostatic forces, since a positively and a negatively charged atom will attract, but this attraction will not last, and eventually will cause repulsion, if a change of pH modifies the ionization status of one of the partners.

Let us now focus not on single atoms, but on the main elements of secondary structures, helices and strands, considered as rigid objects. In Fig. 3 the models of two globular proteins are represented using a very simplified description of their 3D structure: each element of secondary structure, α-helix or β-strand, is approximated by a linear segment, and other elements not included in one of the two previous classes, like turns or coils, totally neglected. Despite the fact that the organization of these elements does not present geodesic symmetry, the comparison with an idealized representation of a tensegret suggests some analogy: secondary structure elements again are held together by attractive forces, counterbalanced by repulsive ones, as in the case of tensegrets. Moreover, let us imagine a denaturation process obtained through the lowering of pH. When some negatively charged group becomes neutral, an attractive force is neutralized. This is

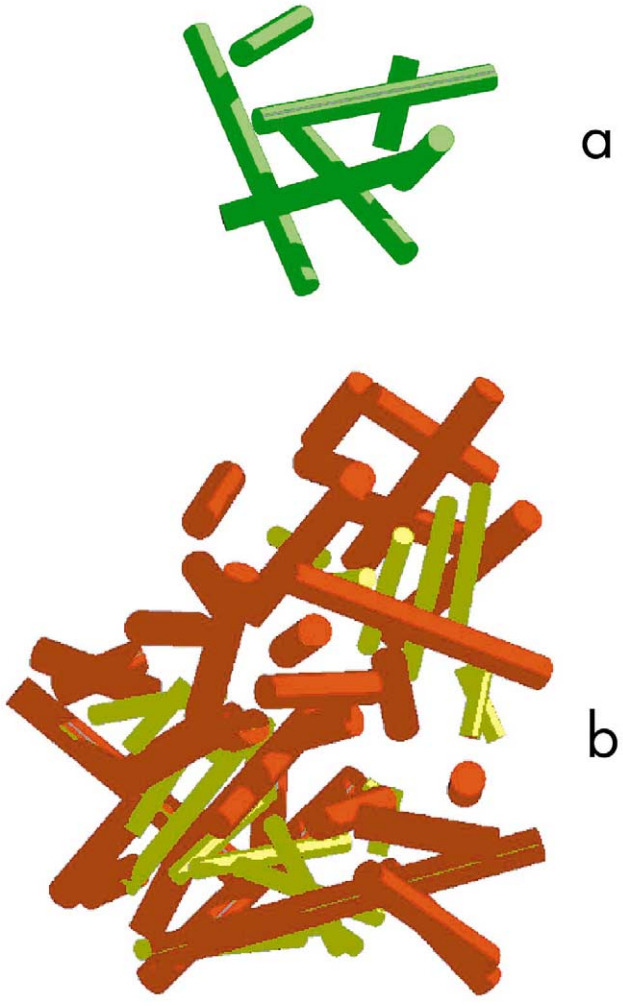

Fig. 3. Simplified representation of (a) myoglobin (coordinates from PDB set 1A6M [26]) and (b) glycogen phosphorylase b (coordinates from PDB set 1GPB [27]). Secondary structure elements, helices and strands, are represented as cylinders (red for helices, yellow for strands), whilst amino acids not belonging to one of the previous secondary structure elements are neglected. This organization must be compared with that of a micelle (Fig. 2c) or with that of an ideal tensegret (Fig. 1). In the latter case, elastic tendons are substituted by attractive or repulsive van der Waals or electrostatic forces.

similar to cutting a cable in a physical tensegret, with the consequence that the entire object has to reassemble, in such a way that the equilibrium between different forces is re-established. This strongly resembles the behavior of tensegrets, where the increase of tension in one part induces limited structural modifications that, increasing and decreasing forces in other parts, maintain the stability of the object.

Although secondary structural elements are not arranged with geodesic symmetry in a single protein molecule, this is partially explained by the fact that proteins are made by a polypeptide chain[1] and secondary structure elements are not organized in space in a fully independent way. A second point is that the length of a polypeptide chain in a protein domain is relatively limited, and so is the number of secondary structure elements. We can say that globular proteins behave like quasi-tensegrets, i.e. they are in some way 'limited' by external constraints that make them deviate from the principles of ideal tensegrity.

If the previous hypothesis is true, a relevant consequence derives from it: prediction of the 3D structure of a protein should be possible, based on a simple scheme of attractive/repulsive forces among secondary structure elements, once the exact location of secondary structure elements is known. This has recently been tested [24] for a group of 12 small proteins: learning the interaction potentials from a set of three known proteins and using a simplified representation of the amino acid side chains, the correct fold could be obtained for most of the test set, only assuming an a priori knowledge of the protein secondary structure.

Moreover, the tensegrity model supports the idea that the different levels of organization of proteins in secondary and tertiary structures have a hierarchical meaning, as the physical result of a two-step folding process. Finally, it simplifies the description of the stability of globular proteins and of the denaturation mechanism.

Acknowledgements: We are indebted to Amos Maritan and Flavio Seno for disclosing to us their results before publication.

References

[1] Connelly, R. and Back, A. (1998) Am. Sci. 86, 142–151.
[2] Fuller, R.B. (1975) Synergetics, Explorations in Geometry of Thinking, Collier Macmillan, London.
[3] Chen, C.S. and Ingher, D.E. (1999) Osteoarthritis Cartilage 7, 81–91.
[4] Wang, J.H.C. (2000) J. Theor. Biol. 202, 33–41.
[5] Stamenovic, D. and Wang, N. (2000) J. Appl. Physiol. 89, 2085–2090.
[6] Heidemann, S.R., Lamoureaux, P. and Bauxbaum, R.E. (2000) J. Appl. Physiol. 89, 1670–1674.
[7] Wang, N., Naruse, K., Stamenovic, D., Fredberg, I.J., Mijailovich, S.M., Toric-Norrelykke, I.M. and Ingher, D.E. (2001) Proc. Natl. Acad. Sci. USA 98, 7765–7770.
[8] Ingher, D.E. (2000) BioEssays 22, 1160–1170.
[9] Howard, J.B., McKinnon, J.T., Makarovsky, Y., Lafleur, A.L. and Johnson, M.E. (1991) Nature 352, 139–141.
[10] Caspar, D.L.D. (1980) Biophys. J. 32, 103–138.
[11] Butcher Jr., J.A. and Grady, W.L. (1984) J. Am. Chem. Soc. 106, 1217–1220.
[12] Berman, H.M., Battistuz, T., Bhat, T.N., Bluhm, W.F., Bourne, P.E., Burkhardt, K., Feng, Z., Gilliland, G.L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padila, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, N. and Zardecki, C. (2002) Acta Crystallogr. D58, 899–907.
[13] Lesk, A.M. (1991) Protein Architecture: A Practical Approach, Oxford University Press, Oxford.
[14] Fischer, D., Tsai, C.-J., Nussinov, R. and Wolfson, H. (1995) Protein Eng. 8, 981–997.
[15] Chothia, C. (1992) Nature 357, 543–544.
[16] Orengo, C.A., Jones, D.T. and Thornton, J.M. (1994) Nature 372, 631–634.
[17] Denton, M. and Marshall, C. (2001) Nature 410, 417.
[18] Chothia, C., Levitt, M. and Richardson, D. (1977) Proc. Natl. Acad. Sci. USA 74, 4130–4134.
[19] Chothia, C., Levitt, M. and Richardson, D. (1981) J. Mol. Biol. 145, 215–250.
[20] Walther, D., Eisenhaber, F. and Argos, P. (1996) J. Mol. Biol. 255, 536–553.
[21] Murzin, A.G. and Finkelstein, A.V. (1988) J. Mol. Biol. 204, 749–769.
[22] Radford, S.E. (2000) Trends Biochem. Sci. 25, 611–618.
[23] Ptitsyn, O.B. (1987) J. Protein Chem. 6, 273–293.

[1] Here we are considering a subunit or a domain, rather than an entire protein.

[24] Hoang, T.X., Seno, F., Banavar, J.R., Cieplak, M. and Maritan, A. (2003) Proteins: Structure, Functions and Genetics, in press.
[25] Silva, A.M. and Rossmann, M.G. (1985) Acta Crystallogr. B41, 147.
[26] Vojtechovsky, J., Chu, K., Berendzen, J., Sweet, R.M. and Schlichting, I. (1999) Biophys. J. 77, 2153–2174.
[27] Johnson, L.N., Acharya, K.R., Jordan, M.D. and McLaughlin, P.J. (1990) J. Mol. Biol. 211, 645–661.