Topological Chirality of Proteins

Chengzhi Liang and Kurt Mislow*

Contribution from the Department of Chemistry, Princeton University, Princeton, New Jersey 08544

Received January 4, 1994®

Abstract: A few rare instances are known in which conformational restriction on polypeptide folding patterns by disulfide cross-links results in topological chirality. We now show that, once the role played by covalently bound cofactors (prosthetic groups) in conjugated proteins is taken into account, topological chirality is in fact more common than previously realized. Iron–sulfur proteins are examples of native proteins in which covalently bound Fe_4S_4 clusters induce topological chirality even in the absence of disulfide cross-links. Quinoproteins with covalently bound cofactors are now recognized to contain catenated substructures, and thus provide the first example of topological complexity in a native protein. The present study strongly suggests that topological chirality may be of wide occurrence among the diverse classes of conjugated proteins.

The constitutional formulae (primary structures) of proteins are given by molecular graphs.[1] With a few exceptions, to be described below, these graphs are reported to be planar.[2] Hence, because nonplanarity is a necessary (though not sufficient) condition for topological chirality,[3] it would appear that even though native proteins are made up of L-amino acids, and higher-order chiralities are imparted to secondary structures by the convolutions of the polypeptide chains, the great majority of these chemically (and geometrically) chiral molecules are *topologically* achiral (Figure 1).

Knots and links are classic examples of topologically nonplanar and chiral objects, yet not a single example has turned up in previous investigations of polypeptide topologies[4-11] of a native protein or polypeptide that contains a knotted or linked (catenated) structural element.[12] This is in stark contrast to the nucleic acids, which exhibit a rich variety of knotted and linked structures, many of them topologically chiral.[13,14] Nonplanarity is also the inevitable result if the molecular graph of a protein contains a $K_{3,3}$ or K_5 subgraph.[2] Yet, until 1993, scorpion variant-3 toxin from *C. sculpturatus* and two closely related polypeptides were the only native proteins known to owe their nonplanarity to such subgraphs: in this family of proteins, in which eight cysteine (or half-cystine) residues form four disulfide bonds, the molecular graph contains $K_{3,3}$ as a subgraph (Figure 2).[15a-c] In 1993 Mao[11] supplied another example of a nonplanar protein with a $K_{3,3}$ subgraph, the light chain of quinoprotein methylamine dehy-

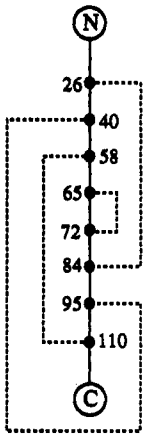

Figure 1. Condensed linear presentation of the molecular graph for Ribonuclease A, an example of a topologically achiral protein. The polypeptide chain is drawn as a vertical line from the N to the C terminals. Cysteine (or half-cystine) residues are numbered and their α-carbons are indicated by solid circles. Intrachain disulfide bonds are shown as dashed lines joining a pair of solid circles.

* Abstract published in *Advance ACS Abstracts*, April 1, 1994.

(1) A *graph* is a set of vertices plus a set of edges that connect some or all of the vertices. Vertices so joined are said to be *adjacent*. In a *molecular graph*, differently labeled (colored) vertices represent different kinds of atoms and differently labeled (colored) edges represent different types of bonds or sequences of bonds, normally covalent. Because the edges in a graph merely symbolize neighborhood relationships between adjacent vertices, the image of a graph is deformable into an infinity of shapes. That is, the edges in a graph can be stretched and bent without limit—so long as they are not severed and rejoined. Thus, a graph is a *topological object*, as distinct from a geometric one.

(2) A graph that can be flattened—embedded in the plane—without the crossing of any edges is said to be *planar;* otherwise it is *nonplanar*. The necessary and sufficient condition for nonplanarity is the presence in the graph of a subgraph that is homeomorphic or contractible to $K_{3,3}$ or K_5. The bipartite graph $K_{3,3}$ consists of two disjoint sets of three vertices each, with each vertex of one set adjacent to all three of the other. The other nonplanar graph, K_5, consists of five vertices that are all adjacent to one another. See: Wilson, R. J. *Introduction to Graph Theory*; Oliver and Boyd: Edinburgh, 1972. In this paper, "a subgraph homeomorphic or contractible to $K_{3,3}$ or K_5" is abbreviated throughout as "a $K_{3,3}$ or K_5 subgraph". Furthermore, all nontrivial knots (like the trefoil knot) and links (like catenated circles) are also topologically nonplanar since, by definition, these objects cannot be projected in the plane without crossings. See: Crowell, R. H.; Fox, R. H. *Introduction to Knot Theory*; Springer-Verlag: New York, 1963.

(3) A graph is *topologically chiral* if it cannot be converted to its mirror image by continuous deformation. Nonplanarity is a prerequisite for topological chirality because a planar graph is achiral in 3-space. It can be proven that topological chirality in a $K_{3,3}$ graph requires a minimum of two non-adjacent colored edges, while topological chirality in a K_5 graph requires a minimum of three colored edges that form an open path (to be published). That is, absent this coloration, $K_{3,3}$ and K_5 graphs are both topologically achiral. Thus, in general, the presence of a $K_{3,3}$ or K_5 subgraph in a molecular graph is a necessary but not a sufficient condition for molecular topological chirality. In the case of the proteins under discussion, however, the vertices of the molecular graphs are all labeled differently because they represent chemically different entities. Under these conditions, molecular graphs containing $K_{3,3}$ or K_5 subgraphs cannot be converted into their topological enantiomorphs by continuous deformation.

(4) (a) Crippen, G. M. *J. Theor. Biol.* **1974**, *45*, 327. (b) Crippen, G. M. *J. Theor. Biol.* **1975**, *51*, 495.

(5) Connolly, M. L.; Kuntz, I. D.; Crippen, G. M. *Biopolymers* **1980**, *19*, 1167.

(6) Klapper, M. H.; Klapper, I. Z. *Biochim. Biophys. Acta* **1980**, *626*, 97.

(7) Thornton, J. M. *J. Mol. Biol.* **1981**, *151*, 261.

(8) (a) Kikuchi, T.; Némethy, G.; Scheraga, H. A. *J. Comput. Chem.* **1986**, *7*, 67. (b) Kikuchi, T.; Némethy, G.; Scheraga, H. A. *J. Comput. Chem.* **1989**, *10*, 287.

(9) Benham, C. J.; Jafri, M. S. *Protein Sci.* **1993**, *2*, 41.

(10) Mao, B. *J. Am. Chem. Soc.* **1989**, *111*, 6132.

(11) Mao, B. *Protein Sci.* **1993**, *2*, 1057.

(12) Because their molecular graphs are planar, the recently reported family of "knotted proteins" (Nguyen, D. L.; Heitz, A.; Chiche, L.; Castro, B.; Boigegrain, R. A.; Favel, A.; Coletti-Previero, M. A. *Biochimie* **1990**, *72*, 431) and "disulfide knot" containing growth factor TGF-β2 (Daopin, S.; Li, M.; Davies, D. R. *PROTEINS: Struct., Funct., Genet.* **1993**, *17*, 176) cannot properly be said to contain knotted structures.[2]

(13) Walba, D. M. *Tetrahedron* **1985**, *41*, 3161.

0002-7863/94/1516-3588$04.50/0