A NOVEL STRUCTURAL MODEL FOR SILK FIBROIN: $\alpha_L\alpha_R\beta$-STRUCTURE

V. I. LIM and S. V. STEINBERG
Institute of Protein Research, USSR Academy of Sciences, Poustchino, Moscow Region, USSR

Received 25 June 1981

1. Introduction

Here we suggest a new type of packing of polypeptide chains in the silk fibroin of *Bombyx mori*. This packing differs completely from that described in [1,2]. It also differs markedly from all known types of structures which occur in globular and fibrous proteins. In our model of silk fibroin, several polypeptide chains form the parallel-chain tape-structure twisted into a flat helix (the tertiary structure) which in turn forms silk fiber (the quaternary structure).

2. Tertiary structure

2.1. $\alpha_L\alpha_R$-Structure

Pauling and Corey's classic works [3,4] give a description of several structures of the pleated-sheet type. Fig.1a presents one of them. This structure is a hydrogen-bonded layer structure of polypeptide chains with all chains similarly oriented. Such a structure will be termed here an $\alpha_L\alpha_R$-structure. Let us consider the $\alpha_L\alpha_R$-structure in which all left-handed α-helical positions are occupied by glycine residues. Such an $\alpha_L\alpha_R$-structure can be formed from polypeptide chains having the amino acid sequence Gly—A—Gly—B—Gly—C, where A,B,C, are any residues except proline. Glycine residues occupying each second position of the chain will be located on one side of the $\alpha_L\alpha_R$-structure and residues A,B,C, on the other. A regular alternation of ridges and grooves on both surfaces allows a good contact between two $\alpha_L\alpha_R$-structures according to the 'ridges into grooves' principle (see fig.1b). Analysis with CPK models shows that the best contact is ensured by glycine surfaces. The geometry of glycine surfaces is so perfect that they can be in contact without any cavities between them.

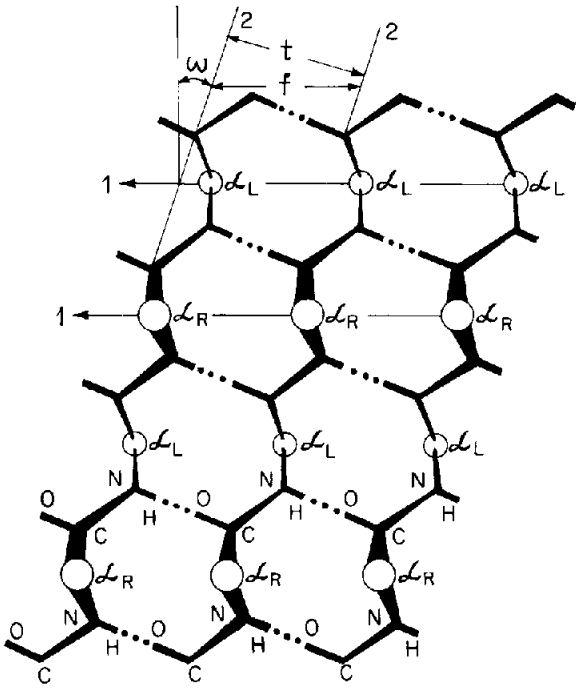


Fig.1. $\alpha_L\alpha_R$-Structure: (a) Circles with symbols α_R are side chains facing the reader. Residues of these side chains are in the region of the right-handed α-helical conformation. Circles with symbols α_L are side chains located on the opposite side of the pleated sheet. Their residues are in the region of the left-handed α-helical conformation. Points denote inter-peptide hydrogen bonds. Both surfaces of an $\alpha_L\alpha_R$-structure represent regularly alternating ridges and grooves. Lines of ridges and grooves (1) form with line (2), parallel to the polypeptide chain, an angle of $90°-\omega$ where $\omega \approx 7°$. The distance between neighbouring chains (t) is 4.7 Å and the distance between neighbouring residues along ridges and grooves (f) is 4.8 Å. Free CO-groups are located on one border of an $\alpha_L\alpha_R$-structure and free NH-groups are located on the other border.

Published by Elsevier/North-Holland Biomedical Press
00145793/81/0000 0000/$02.50

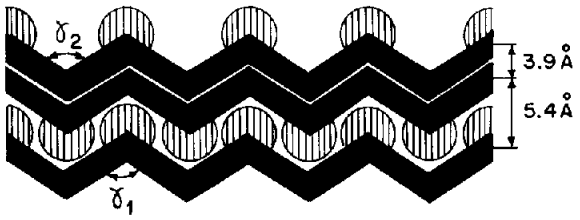


Fig.1. $\alpha_L\alpha_R$-Structure: (b) 'Ridges into grooves' interaction between $\alpha_L\alpha_R$-structures (black zigzag bold lines). View along the lines of ridges and grooves. $\alpha_L\alpha_R$-Structures are formed from polypeptide chains $(Gly-Ala)_n$. Vertical shading denotes methyl groups of alanines. These groups are located on the peaks of ridges. Upper and central $\alpha_L\alpha_R$-structures, separated by 3.9 Å, interact by glycine surfaces while central and lower ones, separated by 5.4 Å, interact by alanine surfaces. Upper and lower $\alpha_L\alpha_R$-structures are strictly parallel to each other. The angles that a central structure forms with each of them are $+2\omega$ and -2ω (for ω see fig.1a). $\gamma_1 = \gamma_2 \simeq 100°$.

2.2. $\alpha_L\alpha_R\beta$-Structure

The polypeptide chain of the silk fibroin of *Bombyx mori* represents a block-copolypeptide in which so-called C_p-fragments alternate with C_s-fragments [5]. C_p- and C_s-fragments account for 60% and 40% (by wt) of the silk fibroin molecule, respectively. The primary structure of a C_p-fragment is known [5,6]: Gly–Ala–Gly–Ala–Gly–Ser–Ala–Ala–Gly–[Ser–Gly–$(Ala–Gly)_2]_8$–Tyr. In this sequence there is a long section [Ser–Gly–$(Ala–Gly)_2]_8$ in which each second position is occupied by a glycine residue. The primary structure of a C_s-fragment is unknown, but it has been reported [7] that glycine residues occupy each second position in a considerable length of its chain. These experimental data on the primary structure enabled us to assemble from several polypeptide chains the parallel-chain tape-structure represented in fig.2. This tape-structure, in turn, can be twisted into a regular flat helix (see fig.3) which we suggest as the tertiary structure for polypeptide chains of silk fibroin. The $\alpha_L\alpha_R$-structures of C_p-fragments are located on one side of this flat helix and the $\alpha_L\alpha_R$-structures of C_s-fragments on the other (see fig.2). On each side of the flat helix $\alpha_L\alpha_R$-structures are packed in parallel. Such a packing allows CO and NH-groups located at the borders of $\alpha_L\alpha_R$-structures (see fig.1a,3) to form

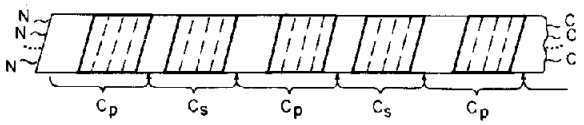


Fig.2. Parallel-chain tape-structure made up of several polypeptide chains of silk fibroin. Parallelograms with dashed lines are $\alpha_L\alpha_R$-structures of equal dimensions made up of fragments in which each second position is occupied by a glycine residue. Glycine surfaces of $\alpha_L\alpha_R$-structures are towards the reader. Dashed lines represent ridges on glycine surfaces. In inter-$\alpha_L\alpha_R$-structural zones the polypeptide chains can have any extended conformation.

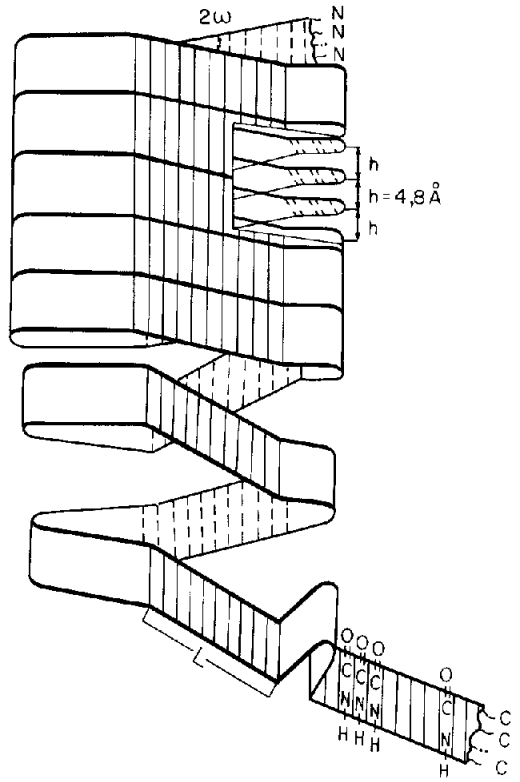


Fig.3. Flat-helical structure (an $\alpha_L\alpha_R\beta$-structure): Vertically shaded regions of the tape are $\alpha_L\alpha_R$-structures. Vertical dashed lines are ridges on the glycine surface of $\alpha_L\alpha_R$-structures, vertical solid lines are ridges on the opposite side of $\alpha_L\alpha_R$-structures. L is the length, in Å, of an $\alpha_L\alpha_R$-structure. The packing of individual polypeptide chains in the flat-helical structure is shown in the upper part of the figure. In inter-$\alpha_L\alpha_R$-structural regions of the tape, each chain forms a β-structural hairpin. Oblique dashed lines are inter-peptide hydrogen bonds in β-hairpins. 'h' is the distance between neighbouring β-hairpins. In the formed part of the flat helix, $\alpha_L\alpha_R$-structures lie in the plane of the figure while the planes of β-hairpins are perpendicular to the plane of the figure. 2ω is the angle between two $\alpha_L\alpha_R$-structures whose glycine surfaces interact with each other according to the 'ridges into grooves' principle. For ω see fig.1a. The number 'm' of polypeptide chains in an $\alpha_L\alpha_R\beta$-structure $= L \cdot \sin 2\omega/4.7$ where 4.7 Å is the distance between chains in an $\alpha_L\alpha_R$-structure (see t in fig.1).

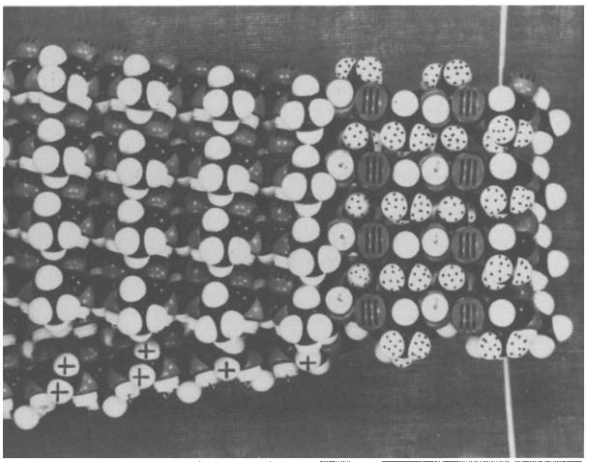
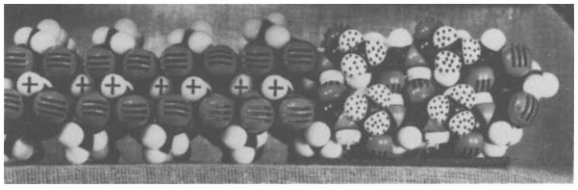


Fig.4. Fragment of an $\alpha_L\alpha_R\beta$-structure assembled from CPK atomic models. (a) View in the direction perpendicular to the glycine bilayer surface. The right part of the figure represents polyalanine β-hairpins. Planes of β-hairpins are perpendicular to the plane of the figure. Methyl groups of alanines are dotted. The left part of the figure represents a fragment of the glycine bilayer. The alanine surface of the upper layer and a small region of the glycine surface (crosses) of the lower layer are towards the reader. The length of irregular regions connecting the ends of a β-hairpin with the glycine bilayer is equal, in this case, to the length of one residue.



(b) View along the ridges of the glycine bilayer. Glycine residues located on the internal surfaces of the glycine bilayer are marked with crosses. β-Hairpins are in the plane of the figure.

interpeptide hydrogen bonds. As a result, a new wide $\alpha_L\alpha_R$-structure consisting of the $\alpha_L\alpha_R$-structures of C_p-fragments is formed on one side of the flat helix and a new wide $\alpha_L\alpha_R$-structure consisting of the $\alpha_L\alpha_R$-structures of C_s-fragments on the other. Both new wide $\alpha_L\alpha_R$-structures interact by their glycine surfaces according to the 'ridges into grooves' principle forming an angle 2ω (see fig.3). This central part of the flat helix will be called a glycine bilayer.

Regions of the tape-structure connecting neighbouring $\alpha_L\alpha_R$-structures of C_p- and C_s-fragments (see fig.2) are located in the flat helix at the borders of the glycine bilayer. The β-structural hairpin is the optimal structure for the polypeptide chain fragments of these regions. Therefore the flat helix in fig.3 was called an $\alpha_L\alpha_R\beta$-structure. In such a structure, the planes of β-structural hairpins are perpendicular to ridges and grooves of the $\alpha_L\alpha_R$-structures.

Hairpins can be sterically joined to the borders of the glycine bilayer provided that the distances 'h' (see fig.3) between neighbouring β-hairpins do not exceed the distance (4.8 Å) between neighbouring residues in a ridge (see 'f' in fig.1a). This condition is satisfied in the case of polyalanine β-hairpins having normal [8,9] van der Waals' radii of atoms (see fig.4a). Such a small distance between β-hairpins is possible because the C^α–C^β-bond of methyl groups is perpendicular to the plane of a neighbouring β-hairpin (see fig.4b). In the C^α–C^β-bond direction, the van der Waals' radius of a methyl group is 1.5–1.7 Å. This range corresponds to the hydrogen radius equal to 1–1.2 Å. When alanine is replaced by a larger residue, the value 'h' can be kept unchanged by removing C^γ, C^δ . . . atoms of a side chain from the interhairpin space. Two bulky side chains of residues occupying neighbouring positions along the chain cannot be simultaneously removed. Therefore dipeptides and longer fragments having only residues with bulky side chains can form only a bend of a β-hairpin and irregular regions joining the ends of a β-hairpin with the glycine bilayer. Thus, in the $\alpha_L\alpha_R\beta$-structure there are definite constraints on the primary structure of the regions forming the glycine bilayer and β-hairpins. The glycine bilayer is formed by fragments in which each second position is occupied by a glycine residue; in β-hairpins, dipeptides with two bulky side chains are forbidden. There are no constraints on β-bends and irregular regions joining the ends of a β-hairpin with the glycine bilayer.

3. Quaternary structure

The best way for the interaction between $\alpha_L\alpha_R\beta$-structures is the packing of their glycine bilayers according to the 'ridges into grooves' principle and formation of inter-$\alpha_L\alpha_R\beta$-structural hydrogen bonds by peptide groups of β-hairpins. Such hydrogen bonds are possible when bulky side chains are absent at the borders of β-hairpins. A 'ridges into grooves' interaction between glycine bilayers is possible when side chains situated on their surfaces do not differ markedly in size. It seems that, in case of silk fibroin, only one side of the glycine bilayer and one border of β-hairpins can participate in such interactions as C_s-frag-

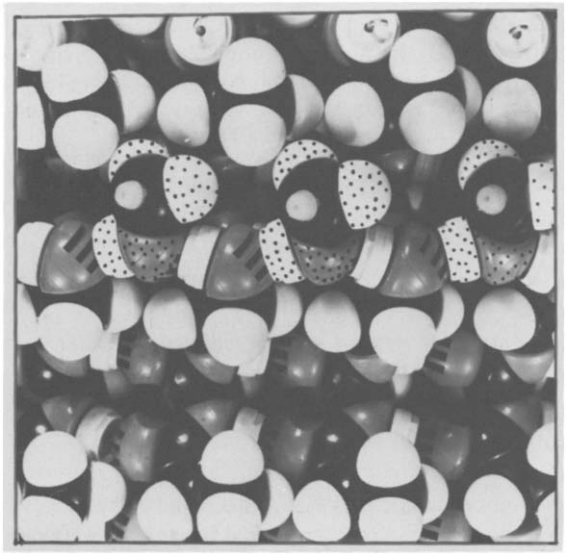


Fig.5. Hydrogen bonds formed by two serine ridges. Dotted serine side chains pertain to one glycine bilayer, those without dots, to the other. Each OH-group of one ridge forms two hydrogen bonds with OH-groups of the other one.

ments of silk fibroin contain, along with a large amount of glycine, alanine and serine residues, a small amount of residues with bulky side chains [5,7]. It should be noted that the interaction between alanine surfaces of $\alpha_L\alpha_R$-structures can be considerably strengthened by replacing each third alanine ridge by a serine one. Namely, such alanine and serine ridges can be formed on one side of the glycine bilayer from the section $[Ser–Gly–(Ala–Gly)_2]_8$ of a C_p-fragment. In this case serine ridges are sterically able to form intermolecular hydrogen bonds (see fig.5) without any changes in the packing of alanine ridges. It is interesting to note that hydrogen bonding between serine ridges is sterically impossible on the surface of either the parallel or antiparallel β-structure.

The above-mentioned features of an $\alpha_L\alpha_R\beta$-structure and of the primary structure of silk fibroin allow us to suggest the layer structure for the silk fibroin fiber (see fig.6a). In each layer of this structure, $\alpha_L\alpha_R\beta$-structures interact with each other only by bend regions of β-hairpins. In turn, each layer along the fiber axis is located antiparallel to the two neighbouring layers. With one neighbouring layer it forms intermolecular hydrogen bonds by peptide groups of β-hairpins, while with the other neighbour it interacts by alanine–serine surfaces of glycine bilayers according to the principle 'ridges into grooves' forming serine hydrogen bonds. Serine hydrogen bonds together with inter-β-hairpin hydrogen bonds hold together all the layers. Each layer also forms with both neighbouring layers zones S_1, S_2 which are loosely packed. This loose packing is due to the presence of a small amount of bulky side chains on the alanine–serine surfaces. Fig.6b presents the layer structure which should result, according to our model, from stretching of silk fibroin along the fiber axis. The fiber is elongated due to elastic stretching of irregular regions joining the ends of β-hairpins with the glycine bilayer and elastic compression of loose regions S_1 and S_2.

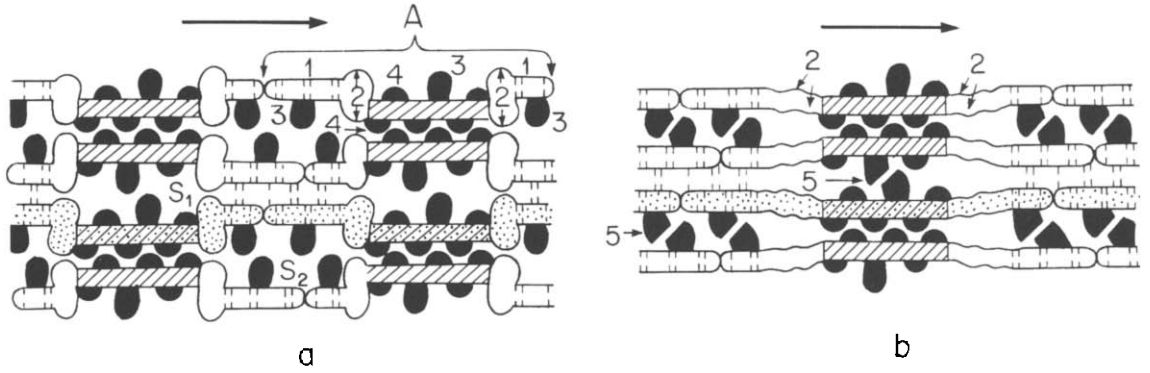


Fig.6. Layer packing of $\alpha_L\alpha_R\beta$-structures for silk fibroin. (a) Unstretched form: *A* is a single $\alpha_L\alpha_R\beta$-structure. Oblique shadowing denotes a glycine bilayer. View along ridges and grooves of glycine bilayers. Identical end faces of flat-helical structures are towards the reader: 1, β-hairpins; 2, irregular regions connecting β-hairpins with the glycine bilayer; 3, bulky side chains; 4, side chains of alanine and serine residues. The dotted region is a single layer. Dashed lines represent inter- and intra-hairpin hydrogen bonds formed by peptide groups. S_1 and S_2 are zones of loose packing. Bold arrow indicates the direction of the fiber axis which is parallel to β-hairpins. (b) Stretched form: 5, schematically represented overlapping of van der Waals' radii and deformation of bulky side chains standard geometry.

4. Conclusion

The spacings observed in our structure do not contradict with those obtained from X-ray patterns of the silk fibroin of *Bombyx mori* [2,10]. The results of infrared dichroism measurements [10] suggest that hydrogen bonds in silk fibroin are oriented approximately perpendicular to the fiber axis. In our fiber structure the main bulk of hydrogen bonds are thus oriented. It is known [11] that the degree of extension of fibers of different silk fibroins is proportional to the content of residues with bulky side chains. This regularity is easily explained within the framework of our model. Extension of a fiber is largely determined by the length of fragments '2' (see fig.6). In turn, the length of these fragments is proportional to the content of bulky side chains in a silk fibroin.

Acknowledgements

The authors thank Drs N. S. Andreeva, N. G. Esipova, N. S. Goel, Yu. A. Lazarev, A. S. Spirin and V. G. Tumanyan for valuable suggestions and critical remarks.

References

[1] Warwicker, J. O. (1954) Acta Cryst. 7, 565–579.
[2] Marsh, R. E., Corey, R. B. and Pauling, L. (1955) Biochim. Biophys. Acta 16, 1–34.
[3] Pauling, L. and Corey, R. B. (1951) Proc. Natl. Acad. Sci. USA 37, 251–256.
[4] Pauling, L. and Corey, R. B. (1951) Proc. Natl. Acad. Sci. USA 37, 729–740.
[5] Lucas, F., Shaw, J. T. B. and Smith, S. G. (1957) Biochem. J. 66, 468–482.
[6] Strydom, D. J., Haylett, T. and Stead, R. H. (1977) Biochem. Biophys. Res. Commun. 79, 932–938.
[7] Lucas, F., Rudall, K. M. (1968) in: Comprehensive Biochemistry (Florkin, M. and Stotz, E. H. eds) vol. 26, pp. 475–558, Elsevier/North-Holland, Amsterdam, New York.
[8] Ramachandran, G. N., Ramakrishnan, C. and Sasisekharan, V. (1963) J. Mol. Biol. 7, 95–99.
[9] Leach, S. J., Nemethy, G. and Scheraga, H. A. (1966) Biopolymers 4, 369–407.
[10] Fraser, R. D. B. and MacRae, T. P. (1973) in: Conformation in Fibrous Proteins and Related Synthetic Polypeptides pp. 293–343, Academic Press, London, New York.
[11] Dickerson, R. E. and Geis, I. (1969) The Structure and Action of Proteins, Harper and Row, London, New York.