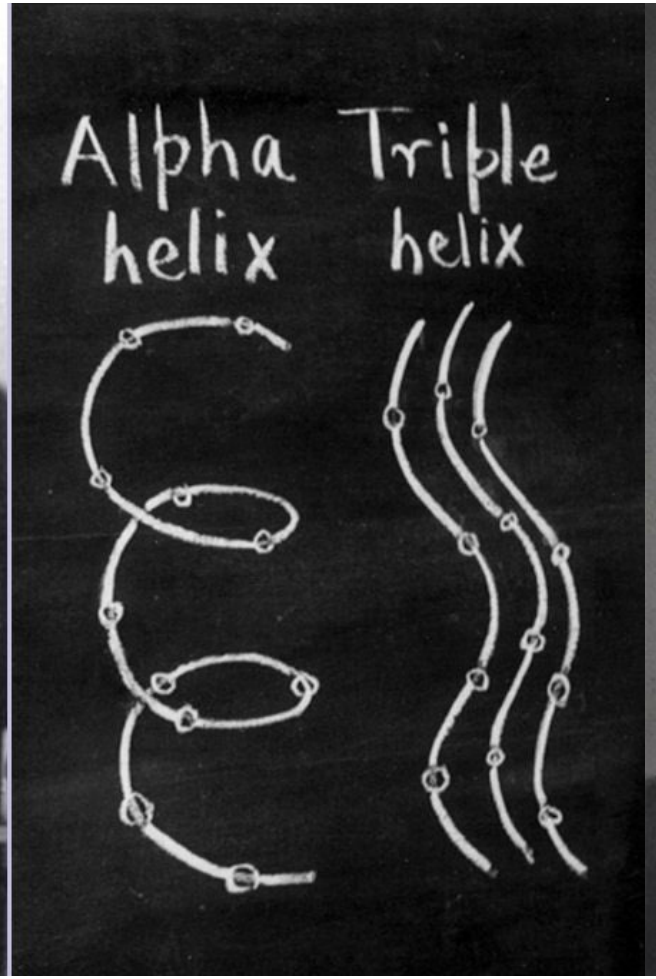
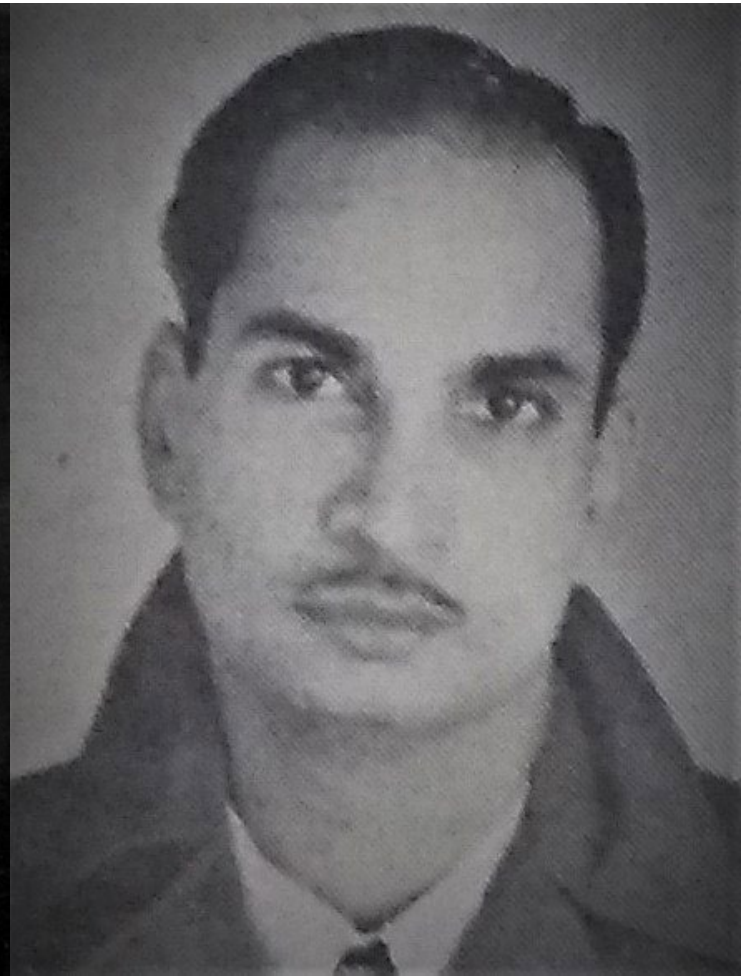
Classic Paper

Structure of Collagen

nature vol 174, August 17, 1954



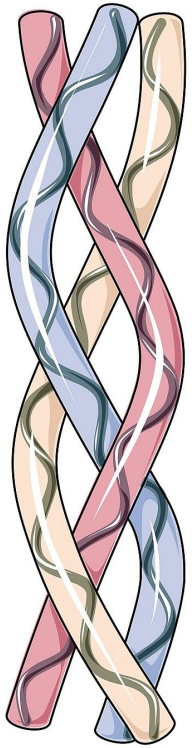


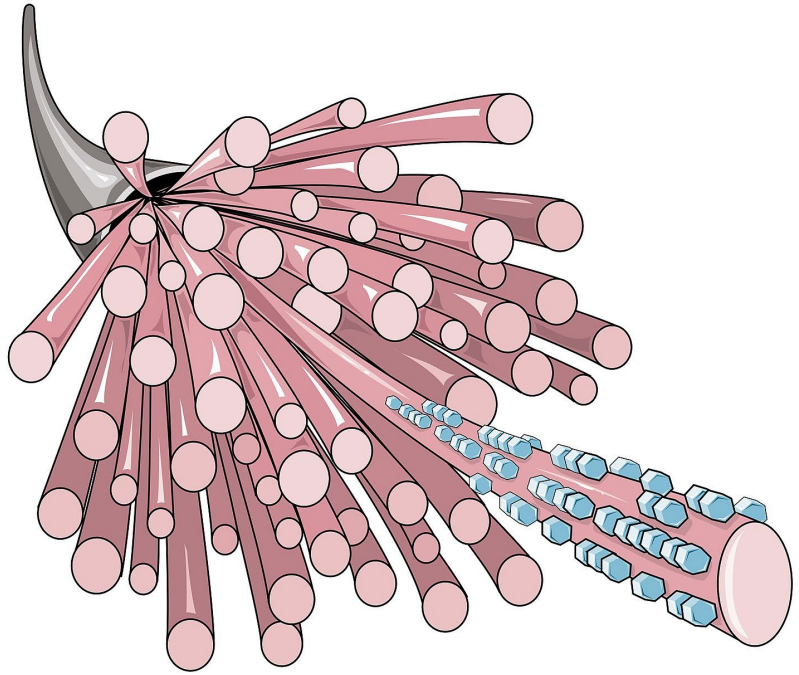
G. N. Ramachandran and Gopinath Kartha

Subrata Bag
12-04-2025

What is collagen ?

Collagen is the main structural protein in the extracellular matrix of the connective tissues of many animals. It is the most abundant protein in mammals, making up 25% to 35% of protein content. Amino acids are bound together to form a triple helix of elongated fibril known as a collagen helix. It is mostly found in cartilage, bones, tendons, ligamen ts, and skin. Vitamin C is vital for collagen synthesis.





A DETAILED X-ray study of collagen fibres obtained from different sources (namely, shark ray, rat tail tendon and kangaroo tail tendon) and a re-examination of the published wide-angle patterns indicate that the unit cell of collagen is hexagonal with $a = 12$–16 A. and $c = 9{\cdot}5$–9 A., the actual values depending on the moisture content. The essential difference of this indexing from those reported earlier[1–3] is that the 2·86-A. meridional arc is here interpreted not as a true meridional reflexion, but as arising from the superposition of two close non-meridional reflexions. A calculation of the angular spread of the arc, using the tilt of the *c*-axis deduced from the spread of the equatorial reflexions, confirms this interpretation. Table 1 shows the good agreement between the calculated and observed spacings of an air-dried and a wet specimen from kangaroo tail tendon.

Table 1

Indices	Dry $a = 13{\cdot}3$, $c = 9{\cdot}55$ A. obs.	calc.	Wet $a = 15{\cdot}0$, $c = 9{\cdot}20$ A. obs.	calc.
100	11·4	11·5	12·9	13·0
200	5·75	5·75	6·5	6·5
210	4·3	4·34	4·8	4·92
101	7·4	7·35	7·6	7·50
201	4·8	4·92	—	4·46
112	3·9	3·87	3·95	3·93
212	3·3	3·22	3·4	3·36
113	2·86	2·87	2·82	2·84
001	—	9·55	9·2	9·20

A structure has been obtained (Fig. 1) which fits the above unit cell and which appears to be in good agreement with infra-red, X-ray and chemical data for collagen. It consists of nine amino-acid residues per unit cell, which corresponds to the observed density. These are linked together to form cylindrical rods, which occur in a hexagonal array. All the residues have the *trans* configuration, and the latest values of Corey and Pauling[4] for the dimensions of the amide group were used for the calculations. The residues are arranged in the form of three helical chains, each of pitch 9·5 A. ($= c$) and containing three residues per turn, with the symmetry 3_1. The three helixes are also arranged with a 3_1 symmetry about the c-axis, and they are held together by means of hydrogen bonds to form the cylindrical rods. Two of the three NH groups in each turn of a chain are linked by hydrogen bonds to an oxygen of each of the other two chains, the NH . . . O distance being 2·80 A. The third NH group points outward from the cylinder, and the nitrogen atom forms part of a proline ring. Of the three α-carbon atoms per turn, a hydrogen attached to one of them (*R*, Fig. 1) could be replaced by a general *R* group to form an amino-acid residue such as arginine or lysine; another (*P*) takes part in forming the proline ring, while the third (*G*) is in such a position that there is no space for either of its hydrogens to be replaced by any other group, so that it could only form part of a glycine residue. These features of the structure could explain the observed range of amino-acid composition for collagen.

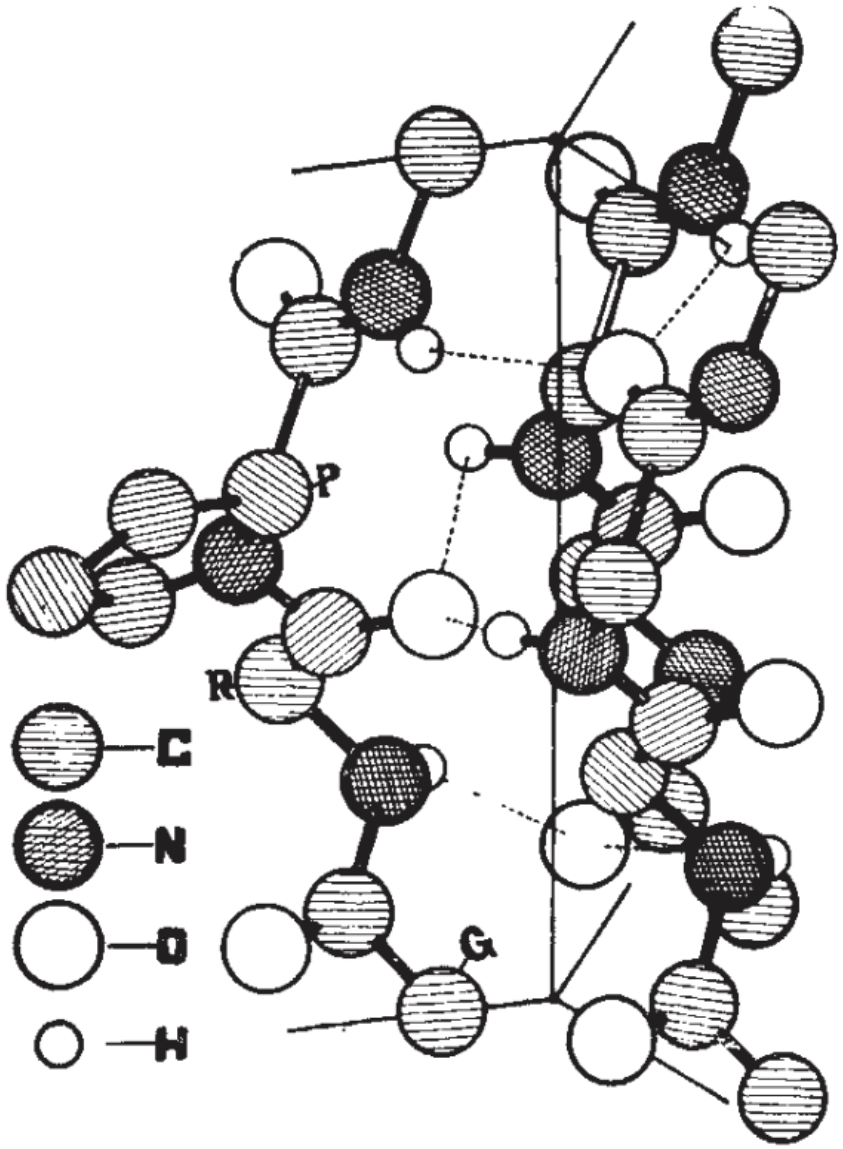


Fig. 1. Diagram showing one of the three-chain cylindrical rods of the structure. The dotted lines indicate the positions of the hydrogen bridges. Only one proline ring is shown, and all hydrogen atoms except those in the NH groups are omitted. The thin lines indicate the directions of the crystal axes

The NH- and CO-bonds are almost exactly perpendicular to the fibre axis, the angle made with the *c*-axis being about 85° in both cases. This agrees with the observed large infra-red dichroism[5]. Structure-factor calculations show fairly good agreement with observation. The individual helixes are unstable, and the stability of the cylindrical rods arises from the hydrogen bonds between the helixes. If these are broken, for example, by heating, the structure would crumble down, which would explain the thermal contraction of collagen. The structure could, however, re-form on cooling, as the chain of amino-acid residues in a single helix need not be ruptured in this process.

It may be mentioned that this structure is essentially different from the three-chain structure of Pauling and Corey[2]. The value $c = 9{\cdot}5$ A. is not critical, and a variation of about 15 per cent is permissible without affecting the main features of the structure.

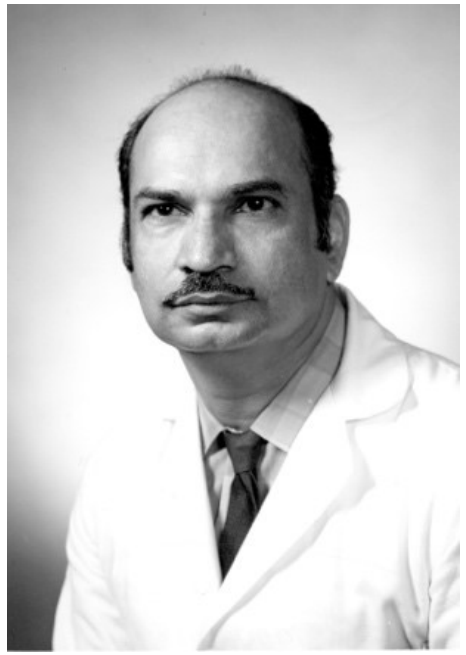
A detailed paper, containing also a discussion of other features of this structure and a critical comparison with previously proposed structures, will be published elsewhere.

Biography

Dr.
G.N. Ramachandran

Dr. G.N. Ramachandran

Born	8 October 1922 Ernakulam, Kingdom of Cochin, British India
Died	7 April 2001 (aged 78) Chennai, Tamil Nadu, India
Nationality	Indian
Alma mater	University of Madras (BSc, MSc, D.Sc) University of Cambridge (PhD)
Known for	Ramachandran plot
	Scientific career
Fields	Biophysics
Institutions	St. Joseph's College, Tiruchirappalli Madras University Indian Institute of Science Cavendish Laboratory
Doctoral advisor	C V Raman



Gopinath Kartha

Born	26 January 1927 Cherthala, Kerala, India
Died	18 June 1984 (aged 57)
Nationality	Indian
Alma mater	University of Madras Andhra University
	Scientific career
Fields	Crystallography
Institutions	Tata Institute of Fundamental Research Cavendish Laboratory National Research Council of Canada Brooklyn Polytechnic Institute Roswell Park Comprehensive Cancer Center
Academic advisors	G.N. Ramachandran