THE POLYPEPTIDE-CHAIN CONFIGURATION IN HEMOGLOBIN AND OTHER GLOBULAR PROTEINS

BY LINUS PAULING AND ROBERT B. COREY

GATES AND CRELLIN LABORATORIES OF CHEMISTRY,* CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA

Communicated March 31, 1951

In the immediately preceding papers we have described several hydrogen-bonded planar-amide configurations of polypeptide chains, and have discussed the evidence bearing on the question of their presence in fibrous proteins. It seems worth while to consider the possibility that these configurations—the pleated sheet, the 3.7-residue α helix, the 5.1-residue γ helix, and the three-chain collagen helix—are represented in molecules of the globular proteins.

It may first be noted that many globular proteins, such as ovalbumin, can on denaturation be converted into a form showing the β-keratin x-ray pattern.[1] The fiber-axis residue distance that is observed, about 3.3 A, is the same as for β keratin, for which we have suggested the pleated-sheet configuration,[2] and it seems reasonable that the same structure should be represented by these denatured proteins. It is, of course, to be expected that a layer structure, such as the pleated sheet, would be assumed by a protein when pressed flat, and the extension of the chains in the pleated-sheet structure makes it reasonable that such a structure should also be assumed by a protein when drawn into a fiber.

The most significant published data bearing on the structure of globular proteins are those on horse carbonmonoxyhemoglobin that have been obtained through the well-planned and diligent efforts of Perutz and his coworkers.[3, 4] These data have been published mainly as a set of sections of a three-dimensional Patterson diagram. We have observed that the data provide some support for the idea that the 3.7-residue helix is a principal feature of the structure of this protein.

Perutz has pointed out that his data indicate that the hemoglobin molecule is about 57 A long, and between 34 A and 57 A in other dimensions, and that there are present rods extending in the 57-A direction, and packed in a pseudohexagonal array, with the centers of the rods about 10.5 A apart. He concluded that the rods probably have the same structure as the molecules in α keratin, for which we have recently suggested the 3.7-residue helical configuration.[5]

There are several facts that favor the view that the 3.7-residue helix is represented in hemoglobin. First, there is the similarity to α keratin, pointed out by Perutz, and the evidence supporting the 3.7-residue helical configuration for the fibrous proteins with the α-keratin structure.[5] Closely related is the fact that from the density and the average residue weight for

hemoglobin one would predict that molecules with this helical configuration would be spaced about 11 A apart (from center to center), in agreement with Perutz's conclusion that the rods in hemoglobin are about 10.5 A apart. (A calculation of this sort at once eliminates the 5.1-residue helix, for which the predicted average spacing of the rods is 14 A.)

Another bit of supporting evidence is provided by the integrated vector density in a strip of the *xz* Patterson section through the origin of the 3-dimensional diagram and in the direction of the axes of the rods. Bragg, Kendrew, and Perutz[6] have reproduced this quantity, plotted as a function of the distance from the origin, in connection with their painstaking analysis

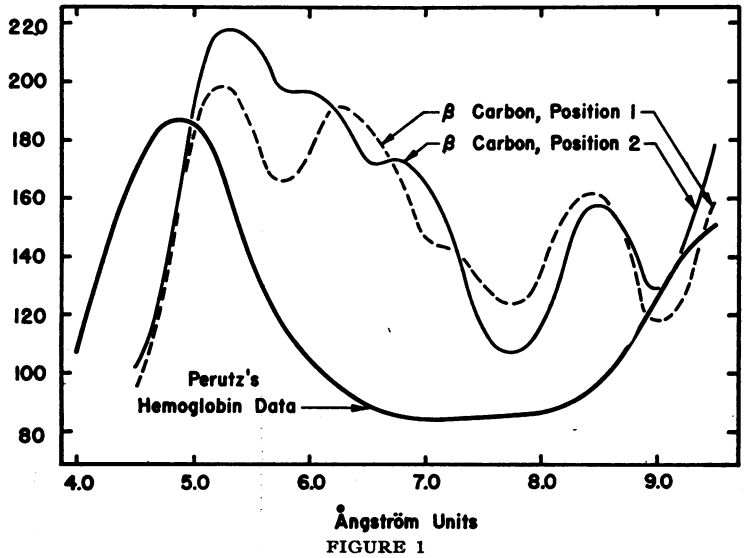

FIGURE 1

Comparison of the radial distribution function calculated for the 5.1-residue helical configuration, with inclusion of a β carbon atom per residue, and the experimental radial distribution function for carbonmonoxyhemoglobin, as calculated from the three-dimensional Patterson function given by Perutz.

of the data for hemoglobin and also for myoglobin[7] and discussion of the correlation of the data with alternative polypeptide configurations. The function has peaks at about 5 A, 11.5 A, 16.5 A, 21.5 A, 27 A, 32 A, etc. We have evaluated a corresponding function for the 3.7-residue helix by including interatomic vectors deviating by not more than 2 A from the direction of the helical axis, and weighting the vectors proportionately to the product of the atomic numbers of the two atoms. The function obtained in this way for an 18-residue 5-turn helix with fiber-axis residue

length 1.53 A has maxima at 5.1 A, 10.6 A, 16.7 A, 21.4 A, 27.5 A, 32.6 A, etc., in excellent agreement with the experimental points.

Another test of the proposed configuration can be made by comparison of the calculated and observed radial distribution functions. Perutz pointed out that the Patterson diagram shows a strong shell at about 5 A from the origin. We have obtained a radial distribution function corresponding to his data for hemoglobin by numerical integration over the contoured Patterson sections published in his paper; this function is shown in figures 1 and 2. It is seen that it has a maximum at about 4.8 A. The calculated radial distribution functions for the 5.1-residue helix are also shown in figure 1. The two curves represent respectively the function for

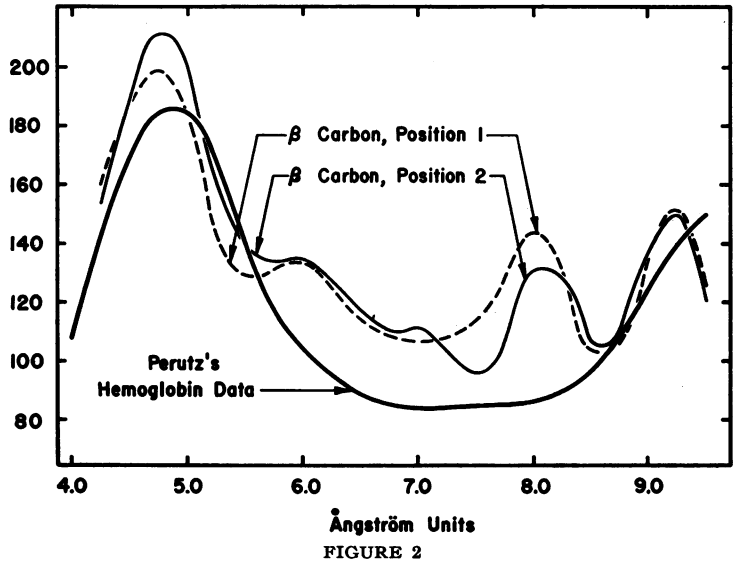

FIGURE 2

Comparison of the radial distribution function calculated for the 3.7-residue helical configuration of the polypeptide chain, with inclusion of a β carbon atom for each residue, and the experimental radial distribution function for hemoglobin.

the four main-chain atoms C, C′, O, and N and a β carbon atom in one of the two alternative positions, and the function for the four main-chain atoms and a β carbon atom in the other position. It is seen that there is no agreement with the hemoglobin curve. The same two calculated radial distribution functions for the 3.7-residue helix are given in figure 2. We think that the rough agreement with the hemoglobin curve is to be considered as significant; it is to be remembered that even with inclusion of the β carbon atom only about 60 per cent of the heavy atoms in the molecule have been taken into consideration in the calculation. The neglected

side-chain atoms are, of course, far more randomly arranged than the main-chain atoms of the helix, and would for this reason tend to distribute their vectors rather uniformly, and thus not to mask the characteristic features of the function due to the main-chain and β carbon atoms.

The comparison of radial distribution functions may thus be construed as giving additional evidence in favor of the suggestion that the rods that Perutz has reported to be present in the hemoglobin molecule have the 3.7-residue helical configuration.

We think that it is not unlikely that this polypeptide configuration is represented in other globular proteins also. In particular, its presence in myoglobin, which is closely related to hemoglobin, would not be surprising; however, it must be pointed out that the Patterson projection for myoglobin on a plane perpendicular to the axis of the rods, given by Bragg, Kendrew, and Perutz,[6] seems hardly to be compatible with this structure. It is possible, of course, that side-chain atoms happen to cooperate effectively in changing the aspect of this projection, or that the axes of the rods do not lie exactly along the direction of projection. The evidence favoring the 3.7-residue helix for myoglobin is contained in Kendrew's description of the myoglobin molecule, as deduced from his data, as consisting of a layer of four rods about 9.5 A apart and with vector maxima spaced 5 A apart in the direction of the axes of the rods. The layers themselves are about 15 A apart, which suggests that if the structure does involve the 3.7-residue helix the side chains are distributed as in crystalline muscle,[5] in which the molecules have an effectively elliptical cross-section, with major and minor diameters 13.1 A and 9.8 A, respectively.

This investigation was aided by grants from The Rockefeller Foundation, The National Foundation for Infantile Paralysis, and The United States Public Health Service. We acknowledge with gratitude the assistance and encouragement of our colleagues in The Gates and Crellin Laboratories of Chemistry throughout the period during which the studies reported in this series of papers and also the investigations on which this work is based were made. We are especially grateful to Professor Verner Schomaker, who has helped by giving us the benefit of both his deep understanding of structural chemistry and his profound critical insight.

* Contribution No. 1556.

[1] Astbury, W. T., and Lomax, R., *J. Chem. Soc.*, **1935**, 846; Astbury, W. T., Dickinson, S., and Bailey, K., *Biochem. J.*, **29**, 2351 (1935).

[2] Pauling, L., and Corey, R. B., these PROCEEDINGS, **37**, 251 (1951).

[3] Boyes-Watson, J., Davidson, E., and Perutz, M. F., *Proc. Roy. Soc.*, **A191**, 83 (1947).

[4] Perutz, M. F., *Ibid.*, **A195**, 474 (1949).

[5] Pauling, L., and Corey, R. B., these PROCEEDINGS, **37**, 261 (1951).

[6] Bragg, W. L., Kendrew, J. C., and Perutz, M. F., *Proc. Roy. Soc.*, **A203**, 321 (1950).

[7] Kendrew, J. C., *Ibid.*, **A201**, 62 (1950).