Stereochemistry of Polypeptide Chain Configurations

Various types of polypeptide chain configurations have been proposed in recent years for proteins as well as polypeptides, e.g. the well-known α-helix for the k-m-e-f proteins, the extended β-chain for silk fibroin, and various structures for simple cyclic peptides having 6 to 10 residues. However, there appears to be no analytical method of writing down these configurations. In connection with the studies on collagen carried out in this laboratory (Ramachandran & Sasisekharan, 1961a,b; Ramachandran, Sasisekharan & Thathachari, 1962), the authors worked out a convenient notation of this type and this is briefly described here. Fuller details are being published elsewhere (Ramachandran, 1962; Sasisekharan, 1962; Ramakrishnan, 1963).

In general, when working out the structure of a long chain polymer, the two types of data to be known are (a) the configuration of the monomer unit and (b) the relative orientations of the linked units at each linkage. In the case of polypeptides,
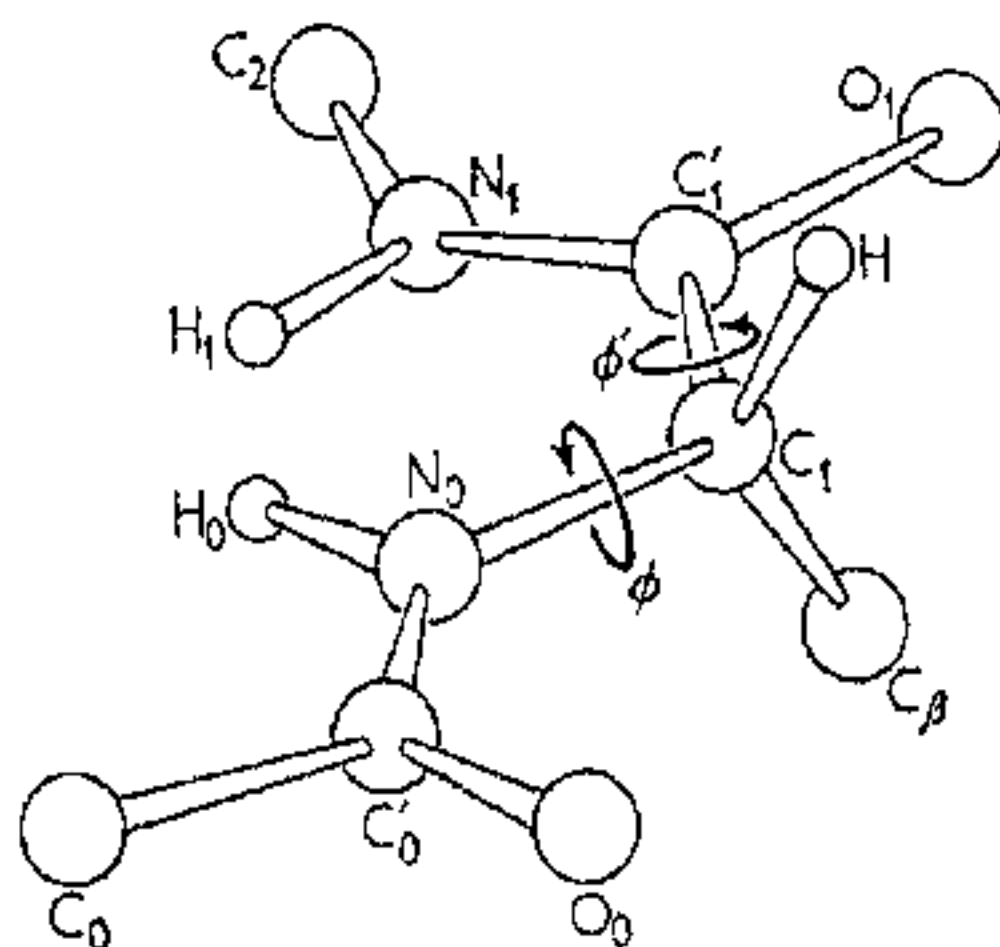

FIG. 1. Standard configuration of two peptide residues linked at an α-carbon atom, in which the planes of both the peptide residues coincide with the plane N_0—αC_1—C_1'. The β-carbon position corresponds to that of residues of L-configuration.

the dimensions of the peptide group itself are very well known, and the generally accepted values are those of Pauling and Corey (PC parameters). In a recent analysis made by the authors, of a number of amino acid and di- and tripeptide structures worked out after the publication of the Pauling–Corey coordinates, it was found that these structures fully confirmed the PC parameters. So also, the planarity of the atoms in the group [αC—$C'O$—NH—αC] is well substantiated, although slight variations of the order of 0·1 Å are observed in certain structures. Consequently, the planar peptide (or amide) group with the PC values was used in all the studies reported below. However, there seems to be a definite probability of the H atoms in the NH group being out of plane.

When two such peptide groups are linked at an α-carbon atom (αC), it is possible to have a rotation of the groups around the bonds N—αC and αC—C'. Consequently the relative configuration of the two groups about the α-carbon atom can be specified by two parameters ϕ and ϕ', which are the angles by which they are rotated about the bonds N—αC and αC—C' respectively from a standard configuration. The conventions we have adopted for the standard configuration ($\phi = \phi' = 0°$) and for the positive

95

sense of rotation for the two angular parameters ϕ and ϕ' are indicated in Fig. 1. When there are a number of peptide groups linked via α-carbon atoms C_1, C_2, etc., the configuration of the chain as a whole can be specified by the parameters (ϕ_1, ϕ_1'), (ϕ_2, ϕ_2'), etc., and the notation is equally valid for a cyclic polypeptide. Mizushima & Shimanouchi (1961) have proposed a similar notation; however, they allow for a rotation which distorts the planar peptide group, and therefore require three angles θ_1, θ_2, θ_3. The relations between their notation and that used in the present investigation are as follows: $\theta_1 = 180°$, $\theta_2 = \phi'$, $\theta_3 = \phi + 180°$.

It is quite clear that all values of ϕ and ϕ' may not be allowed owing to the possible short contacts between the atoms in the adjacent residues. When these are worked out, it is found that there are whole regions of forbidden values for ϕ and ϕ'. However, the demarcation of these regions depends on the choice of the permitted van der Waals contact distances. Two such sets, termed 'normally allowed' and 'outer limit', were worked out from a detailed analysis of the available structural data on various organic compounds, including in particular amino acids and peptides (Table 1).

TABLE 1

Minimum contact distances assumed

Contact	Normally allowed (Å)	Outer limit (Å)
C ... C	3·20	3·00
C ... O	2·80	2·70
C ... N	2·90	2·80
C ... H	2·40	2·20
O ... O	2·80	2·70
O ... N	2·70	2·60
O ... H	2·40	2·20
N ... N	2·70	2·60
N ... H	2·40	2·20
H ... H	2·00	1·90

On the basis of these, the permitted ranges for ϕ and ϕ' shown in Fig. 2 were obtained, corresponding to an angle of 110° between the bonds N_0—αC_1 and αC_1—C_1' at the α-carbon atom. In working these out, all the contacts between the atoms in two neighbouring peptide groups were taken into account, including the β-carbon atoms. If the angle at the α-carbon atom is varied, say from 105° to 115°, then the allowed regions are altered slightly.

The available data on known di- and tripeptide structures and polypeptide chains were examined to see whether their configurations occurred within the allowed ranges of (ϕ, ϕ'). It was in fact found that they all occurred within the outer limits but that, in addition, there was a clustering of the configurations near certain specific types. The main conclusions were the following.

(a) Invariably, whenever the α-carbon concerned belonged to a glycyl residue, i.e. when it had no β-carbon atom of a side-chain attached to it, the configuration was close to $(0°, 180°)$.

(b) On the other hand, when there was a β-carbon atom, the clustering was close to $(120°, 150°)$. In fact, the value of $\phi' \simeq 150°$ was invariably found for the carboxyl group in almost all amino acid structures with a β-carbon atom.

(c) The silk structure, though close to the extended chain, is not fully extended. Its configuration (40°, 140°) is slightly shifted from the pure glycine configuration (0°, 180°) owing to the presence of alanine and other residues.

(d) Both the right- and left-handed α-helices lie within the outer limits but outside the inner limits. Since $\phi \simeq 120°$ for a right-handed α-helix, a proline side-chain can be attached readily to this helix. However, the resulting short contacts between the side-chain atoms and those of the backbone in the helix lead to a break in the helix direction. The hydrogen bonds involving the NH groups of the two succeeding residues are broken, but thereafter the helix can re-form. The angle between the straight helical portions on either side of a proline residue cannot be less than about 35°. In fact, even
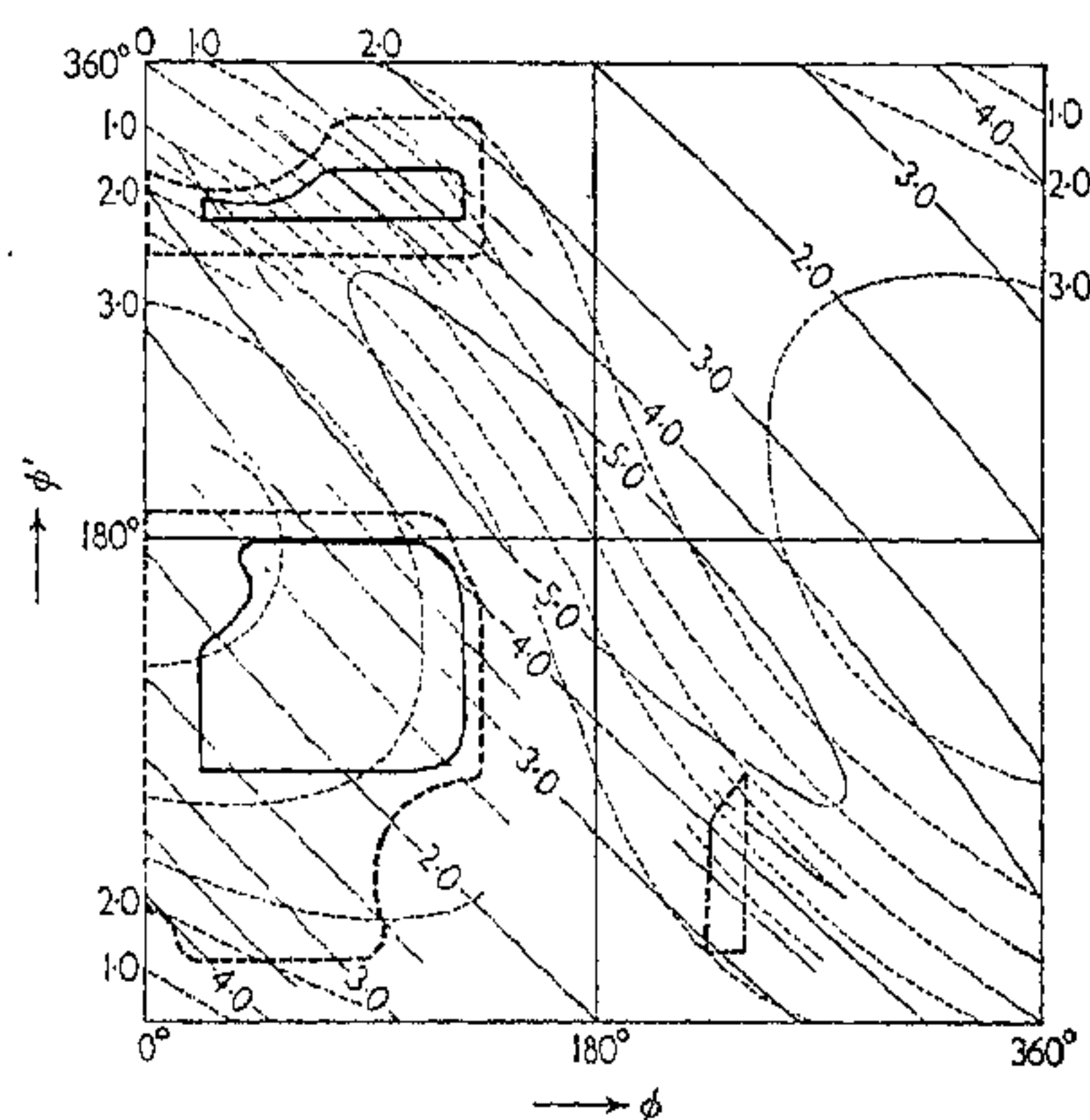

FIG. 2. The fully allowed (————) and outer limit (— — —) regions of (ϕ, ϕ') for angle N—αC—C' = 110° along with the configurations of various known di-, tri- and polypeptide and protein structures.

at 35°, some short contacts occur between the backbone atoms of the peptide residue containing the proline ring and its neighbouring residue, which may, however, be relieved by distorting the residues from their planar configuration.

On the other hand, the angle ϕ is close to 240° for a left-handed α-helix and a proline residue cannot be attached to this helix (in the L-configuration) without very considerable distortion of the helix.

From a careful consideration of the short contacts in α_R and α_L, it is found that right-handed α-helices are more stable than the left-handed ones. The configuration of the π-helix also lies only within the outer limits and it also has every NH hydrogen bonded to a CO as in the α-helix. Hence it is quite likely that the π-helix too will be found in some polypeptides. It is interesting to note that no simple peptide which has been studied so far by X-ray methods has a configuration near that of the α-helix, which corresponds to (130°, 310°). This shows that the α-helical configuration is stable only because of the large number of hydrogen bonds, and that if these were disrupted the chain would immediately collapse.

(e) On the other hand, the configuration of the individual chains in the collagen helix (near 120°, 150°) is a very natural one. It is therefore understandable why three such chains can come together and form a triple chain. So also, the configuration $\phi = 120°$ is demanded for an *imino* acid residue (because of the N being linked to αC *via* the five-membered ring) and the value $\phi' = 150°$ is a natural one for all amino acids, as mentioned earlier. The large proportion of imino acid residues, which is characteristic of collagen, is therefore another factor contributing to the stability of the collagen triple chain.

(f) The right-handed γ-helix lies outside even the outer limits and the two contacts $C_\beta ... C_0'$ and $C_\beta ... O_0$ are very bad, being as short as 2·75 and 2·45 Å. Thus the
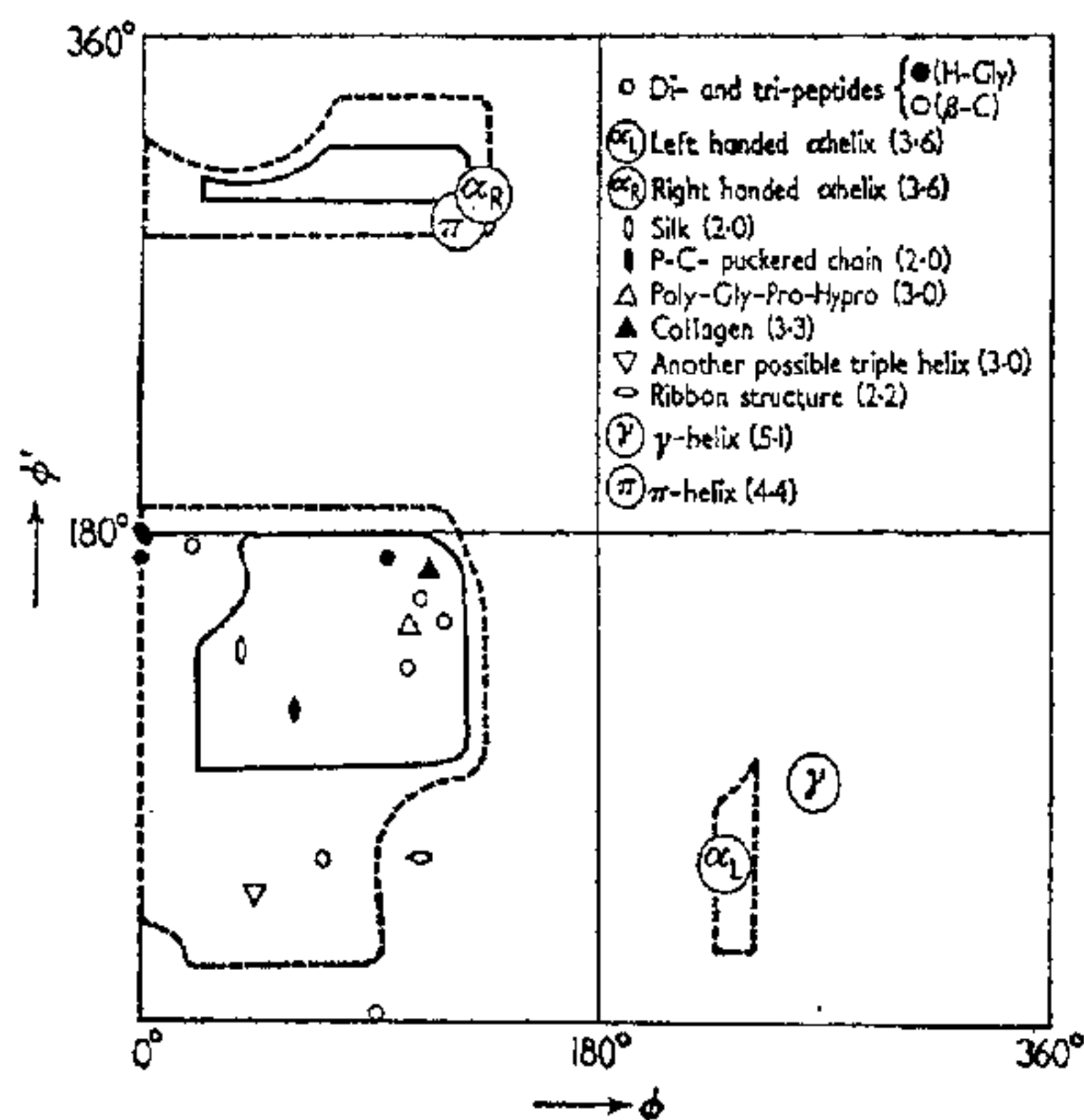

FIG. 3. Contours of constant n (——) and constant h (- - - - - -) corresponding to the angle N—αC—C' = 110°. The boundaries of the fully allowed and outer limit regions are also shown.

γ-helix is unlikely to occur in polypeptide structures. The left-handed γ-helix is equally bad.

(g) In the so-called ribbon structure (2·2₇ helix (Donohue, 1953), see Fig. 2), although the configuration lies just outside the outer limit, the short contact which occurs in it, namely $N_0 ... O_1 = 1·90$ Å, is involved in the hydrogen bond $N_0—H_0 ... O_1$, and is therefore not objectionable. Thus it is likely that this structure will be found in some compounds. Very recent X-ray studies made in Japan (Kakudo, private communication) indicate that this configuration occurs twice in a tetrapeptide structure.

(h) There is another possible triple helix (V of Fig. 2) which is within the outer limit region. Although this cannot accommodate a pyrrolidine ring, it can form inter-chain hydrogen bonds similar to the collagen helix. This structure deserves further study.

Analytical formulae have also been worked out for calculating the number of residues per turn (n) and the residue height (h) of the helical structure which is formed when the same configuration (ϕ, ϕ') is repeated at every α-carbon atom in a polypeptide chain.

The data thus obtained for the angle $N—\alpha C—C' = 110°$ are shown graphically in Fig. 3 along with the fully allowed and outer limit regions. From a study of these curves, it has been possible to show why collagen cannot be stretched by more than a few per cent, although the residue height of near 3 Å is much less than the residue height of 3·6 Å for a fully extended structure. Complete numerical data of n and h corresponding to values of 105°, 110° and 115° for this angle $\overset{\frown}{N\alpha C}C'$ will be published elsewhere (Ramakrishnan, 1963).

It should be mentioned that the above discussions all relate to the natural configuration of L-polypeptides. When D-amino acids are present, the allowed configurations will correspond to $(-\phi, -\phi')$, where (ϕ, ϕ') refers to the L-amino acid. The application of these ideas to small cyclic peptides is likely to be rewarding.

Department of Physics   G. N. RAMACHANDRAN
University of Madras   C. RAMAKRISHNAN
Madras 25, India   V. SASISEKHARAN

Received 27 December 1962

REFERENCES

Donohue, J. (1953). *Proc. Nat. Acad. Sci., Wash.* **39**, 470.

Mizushima, S. & Shimanouchi, T. (1961). *Advanc. in Enzymol.* **23**, 1.

Ramachandran, G. N. (1962). In *Collagen*, ed. by N. Ramanathan, p. 3. New York: John Wiley.

Ramachandran, G. N. & Sasisekharan, V. (1961a). *Nature*, **190**, 1004.

Ramachandran, G. N. & Sasisekharan, V. (1961b). *Curr. Sci.* **30**, 127.

Ramachandran, G. N., Sasisekharan, V. & Thathachari, Y. T. (1962). In *Collagen*, ed. by N. Ramanathan, p. 81. New York: John Wiley.

Ramakrishnan, C. (1963). *Proc. Ind. Acad. Sci.* A, in the press.

Sasisekharan, V. (1962). In *Collagen*, ed. by N. Ramanathan, p. 39. New York: John Wiley.

The Ramachandran Map—A Recent Assessment[1]

The literature on conformational analysis of peptides is enormous, but it would not be far from the truth to say that most of the important questions were answered before 1964. The role of steric factors in determining the allowed values of the dihedral angles about the C_α atoms was largely dealt with by Ramachandran et al.[2] and the role of the peptide dipole in the stabilization of secondary structure was thoroughly treated by Arridge and Cannon[3]. Subsequent conformational analyses have not really changed our overall perspective of what is and what is not allowed in protein secondary structure. The major advances since these early treatments have been in the area of supersecondary arrangements (e.g., Cohen et al., the analysis and prediction of chain topologies and folding patterns (e.g., Sternberg and Thornton[5]), and the establishment of criteria by which the accuracy of predictions may be assessed (e.g., Cohen and Sternberg[6], Nishikawa[7]).

The early conformational analyses were reviewed by Ramachandran and Sasisekharan[8]. Other useful reviews are to be found in the Proceedings of the Madras Conference on the Conformation of Biopolymers (Ramachandran[9]) and in the annual Specialist Periodical Reports of the Royal Society of Chemistry, Amino Acids, Peptides and Proteins.

Ramachandran et al.[2] showed that the possible conformations that could be adopted by small peptides are severely limited by the steric constraints of the atoms actually coming into contact. Most of the energetically favourable regions of the Ramachandran map are predicted by hard-sphere analyses. Small changes in the coordinates of the dipeptide atoms can completely eliminate the possibility of any allowed regions; thus, it is fairly safe to say that local secondary structure is primarily determined by steric interactions. . . . The influence of minor changes in the geometry of the peptide unit was also explored by Gibson and Scheraga[10], who showed that allowing a degree of flexibility in the structure considerably increased the conformational possibilities.

The inclusion of electrostatic interactions (Brant et al.[11]) does not significantly change the allowed conformations derived from steric analysis. There are minor alterations in the shape of the contours and in the positions of the absolute minima on the map, but only a seasoned connoisseur of Ramachandran plots would be able to detect whether he was viewing a potential map that included or excluded electrostatic interactions, because alterations in the parameters of the energy function might cause more significant alterations.

1. Taken from Rogers, N. K. in Prediction of Protein Structure and the Principles of Protein Conformation (ed. Fasman, G. D.), Plenum Press, New York and London, 1989, pp. 373–374.
2. Ramachandran, G. N., Ramakrishnan, C. and Sasisekharan, V., J. Mol. Biol., 1963, 7, 95. (reproduced in this issue).
3. Arridge, R. G. C. and Cannon, C. G., Proc. R. Soc. London, 1964, A278, 91.
4. Cohen, F. E., Sternberg, M. J. E. and Taylor, W. R., Nature, 1980, 285, 378.
5. Sternberg, M. J. E. and Thornton, J. M., Nature, 1978, 271, 15.
6. Cohen, F. E. and Sternberg, M. J. E., J. Mol. Biol., 1980, 138, 321.
7. Nishikawa, K., Biochim. Biophys. Acta, 1983, 748, 285.
8. Ramachandran, G. N. and Sasisekharan, V., Adv. Protein Chem., 1968, 23, 283.
9. Ramachandran, G. N., Conformation of Biopolymers: Proceedings of the 1967 Conference at Madras, Academic Press, New York, 1967.
10. Gibson, K. D. and Scheraga, H. A., Biopolymers, 1966, 4, 709.
11. Brant, D. A., Miller, W. G. and Flory, P. J., J. Mol. Biol., 1967, 23, 47.