SOME POSSIBLE MODIFICATIONS OF THE α-HELIX CONFIGURATION INVOLVING CHANGES OF DIRECTION AND SENSE OF THE HELIX

by

H. LINDLEY*

*Gates and Crellin Laboratories of Chemistry**, California Institute of Technology, Pasadena, Calif. (U.S.A.)*

The α-helix configuration of the peptide chain proposed by PAULING, COREY AND BRANSON[1] is an exceedingly attractive hypothesis of protein structure. It now seems almost certain that synthetic polypeptides such as poly-γ-methyl-L-glutamate and poly-γ-benzyl-L-glutamate can be obtained in a form which has a structure made up of a hexagonal array of α-helices. Structures involving the α-helix are the most convincing so far proposed for fibrous proteins such as wool and hair.

LINDLEY AND ROLLETT[2] have recently proposed a structure for zinc-insulin which is based on the α-helix and which seems to be in general agreement with chemical, physical, and crystallographic evidence. All that can be definitely said of these results so far is that they are compatible with the idea that the α-helix configuration is an important part of globular protein structure.

There are a number of observations which indicate some possible limitations of the α-helix as applied to protein structure and thus suggest the kind of modifications that may be necessary. It has already been shown that a straight α-helix is incompatible with the X-ray patterns of α-keratin and recourse has been had to compound helices or "coiled coils" in which the axis of the α-helix itself is helical[3,4]. X-ray data on the structure of haemoglobin have been interpreted as implying that the α-helix, if present, must be bent several times through 180°. The only suggestion so far put forward to achieve this involves proline residues and it is known that there are insufficient of these in the haemoglobin molecule to account for the extensive bending required[5], even if only one proline residue were required for each 180° bend.

More specific problems are raised by the mode of incorporation of L-proline and L-cystine into the α-helix. L-Proline will not fit into a L.H. (left-handed) α-helix at all, and only fits an undistorted R.H. (right-handed) α-helix when it is one of the first three residues from the end of the chain carrying the free amino group. It is certain from chemical data on proteins that other structural possibilities must exist.

With the elucidation of the complete chemical structure of insulin by SANGER and co-workers two possibilities for the incorporation of L-cystine into proteins must be regarded as proved[6]. These are the interchain disulphide bond and the intrachain

---

* Present address, Wool Textile Laboratories C.S.I.R.O., Parkville N. 2., Victoria, Australia.

** Contribution No. 1979.

disulphide bond in which the two halves of the cystine molecule are separated by comparatively few amino-acid residues. The interchain bond is difficult to reconcile with the elastic properties of wool and hair. Specifically the evidence suggests that α–β transformation which occurs on stretching hair is not accompanied by rupture of (presumed) interchain sulphur bonds. However it would seem probable that conversion of an α-helix structure to the β-form would impose considerable rotational strain on any interchain sulphur bonds unless some very definite structural relationship is assumed such as that all the sulphur bonds are separated by both an integral number of turns of the helix and an odd number of residues. The general problem of intrachain disulphide bonds has been discussed by LINDLEY AND ROLLETT[2], who have pointed out that no such bond involving L-cystine can be formed on an undistorted α-helix.

It has often been assumed, more particularly on chemical evidence, that these difficulties can be explained as due to specific side-chain effects, but no concrete structural possibilities for achieving this have been presented. The purpose of the present paper is to present specific ways of modifying the α-helix configuration by the use of hypothetical amino-acid sequences to illustrate how some of the limitations of the simple α-helix may be circumvented.

1. *180° bend in the α-helix*

Fig. 1 illustrates how a L. H. α-helix might be turned through 180°. The amino-acid sequence involved is –ser–gly–X–X–glu(NH_2)–asp(OH)–gly–. The two glycine residues are necessary because of the close approach of the α-helices at this point, but the bend is not quite planar so that these are the only two positions which

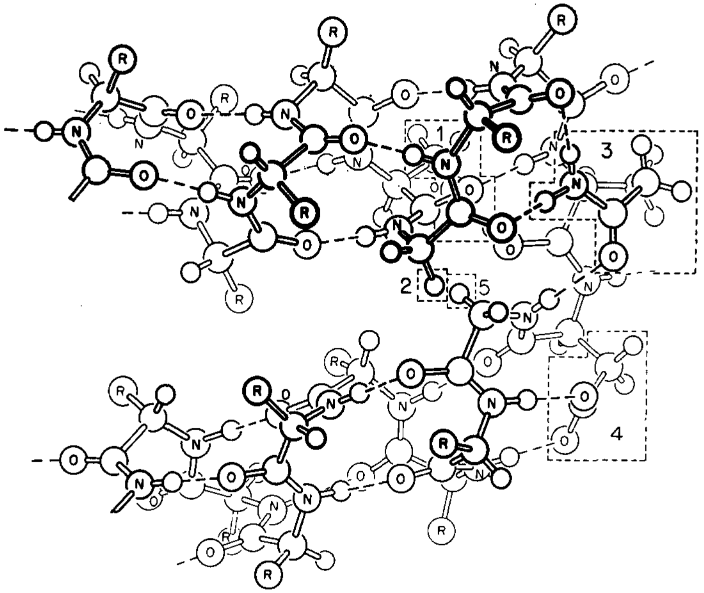

Fig. 1. Diagram of a possible configuration of a 180° bend in a L. H. α-helix involving the amino-acid sequence –ser(1)–gly(2)–X–X–glutamine(3)–aspartic(4)–gly(5)–. The serine side chain (1) forms an OH···O bond with the carbonyl group of the glutamine (3) residue which cannot be shown clearly because of overlying atoms. The close approach of the chains at one point which necessitates the glycine residues (2 and 5) can however be clearly seen.

approach so closely. It is not suggested that this is the only manner in which a 180° bend can occur on an α-helix.

It does illustrate some important points, however. Thus it utilises only amino-acids which occur in relatively high proportion in most proteins and specifically it does not involve L-proline.

This is of importance since there has been a tendency to regard proline as the prime cause of bending of α-helices despite the fact that in the case of haemoglobin there are not sufficient proline residues to allow even one to each bend of the α-helix required from X-ray structural investigations. There is no evidence for the occurrence in haemoglobin of the specific sequence suggested here, but it is of interest that haemoglobin contains more than sufficient glycine, serine, glutamic and aspartic acids for this sequence to account for the number of bends of α-helix which are required (if indeed the α-helix is the basic structure of haemoglobin).

One other aspect of the structure is worthy of mention. It can be thought of as being derived from a simple straight α-helix by the rupture of four main-chain NH···O bonds. However in the final structure there are formed five new NH···O bonds from the glutamine and aspartic side chains and also one OH···O bond from the serine residue. Energetically therefore the configuration would be superior to a simple α-helix, which is not reinforced by side chain H-bonding. This is true for all structures which can be regarded as derived from the α-helix in which more than half of the broken main chain H-bonds can be compensated by forming new H-bonds with reactive side chains, and thus other structures involving less complete H-bonding of main chain carbonyl and imino groups than the one illustrated in Fig. 1 could be energetically favourable compared to the α-helix.

Finally it may be pointed out that this type of configuration is probably only possible with a L.H. α-helix. For a R.H. α-helix the R groups emerge from the helix pointing away from the bend and compensation of the broken main chain H-bonds by side chains is more difficult.

2. *Change of sense of the α-helix from L.H. to R.H. and the intrachain disulphide bond*

The reasons for postulating a change of sense of the α-helix from L.H. to R.H. to accommodate an intrachain disulphide bond have been discussed by LINDLEY AND ROLLETT[2] and their paper includes a diagram of a proposed structure of this kind for the A chain of insulin. If we consider for simplicity a general case similar to that postulated for the A chain of insulin in which residues 1–9 form a L.H. α-helix and residues 9 onward a R.H. helix, then the carbonyl groups of residues 6, 7, and 8 and the imino groups of residues 10, 11, and 12 are not involved in either of the α-helices. Theoretically there are thus nine possible ways of forming one H-bond between the two segments of helix as well as possibilities of forming two H-bonds simultaneously. In fact a number of these theoretical possibilities are eliminated on steric grounds when an attempt is made to construct models of them. For a L.H. to R.H. change there are probably only eight possibilities, *i.e.*, involving 6–10, 6–11, 7–12, 8–10, 8–11, and 8–12 bonds and the additional 6–11 + 8–10 and 7–12 + 8–11 combinations. It is difficult to be dogmatic concerning possible intrachain cystine bonds because the covalent sulphur bond may compensate energetically for a number of H-bonds and moreover possible variations in dihedral angles from 70° to 110° have been reported as well as an unusually long H-bond for N,N'diglycylcystine[7].

Within these limitations only the two intrachain disulphide bonds discussed by Lindley and Rollett[2] have been found to conform to the rigid requirements of bond angles, bond lengths, and van der Waals distances. Other possibilities involving separation of the two halves of the cystine molecule by two, three or possibly five residues could probably occur if one or more of these requirements is relaxed.

Fig. 2 is a diagram of the alternative intrachain disulphide bond which was rejected by Lindley and Rollett[2] as a structure for the A chain of insulin on the grounds that it was less well H-bonded when constructed with the specific amino-acid sequence required. In Fig. 2 a more favourable amino-acid sequence has been chosen and both CO groups are bonded by one arginine residue. Neither of the main chain NH groups can be bonded by any side chain whatsoever. The structure shown has the merit of demonstrating how an arginine residue can be incorporated without the side chain projecting unduly from the helix and can in addition to H-bonding the main chain CO groups also form a very satisfactory intrachain salt link with a suitably placed glutamic acid residue.

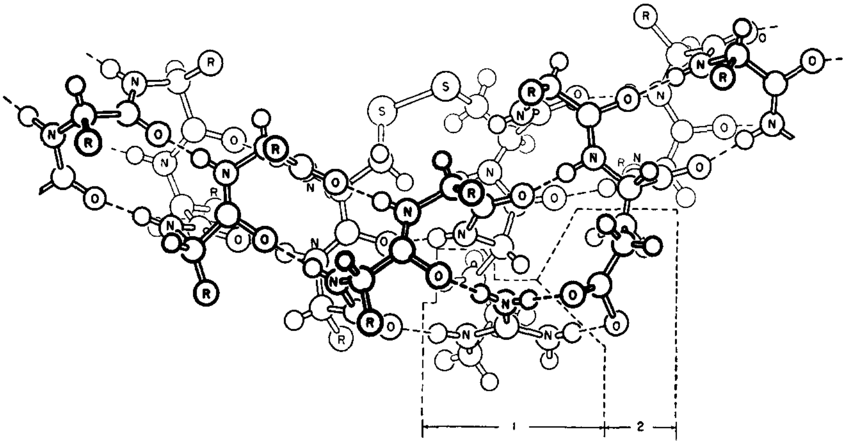

Fig. 2. A possible intrachain disulphide bond involving a change of sense of an α-helix from L. H. to R. H. The amino-acid sequence illustrated is –cys–X–X–X–arg–cys–X–glu–, and shows how the arginine residue may form an intrachain salt link in addition to stabilising the sense change by H-bonding main chain carbonyl groups.

3. *The incorporation of* L-*proline into the α-helix*

L-Proline does not fit into a L. H. α-helix at all and only fits well into an undistorted R. H. α-helix if it is one of the first three residues from the amino-end. It is this fact which has led to various suggestions that the occurrence of proline in an α-helix results in its changing direction. However if one postulates a change of sense of the α-helix from L. H. to R. H. the result is virtually to create a new amino-end of a R. H. helix and proline fits naturally into such a structure. Fig. 3 shows one possible way in which this can be achieved: for this particular sense change the proline might equally well be one residue further along the chain or both positions could be occupied by proline. The structure has been stabilised by using an arginine residue to bond both main chain carbonyl groups, but its use is somewhat different from that in the previous example. In this case the arginine is so situated that it could readily form an interchain salt link. Other possibilities of bonding the CO groups exist and the NH group could also be bonded by a suitably placed side chain. The structure is of interest in that it involves proline (and could in fact involve prolyl-proline) without producing any change in direction of the α-helix. Looked at from this viewpoint it may be that L-proline has two important structural functions in proteins: to bring about a change of sense of an α-helix from L.H. to R.H. and perhaps more specifically to do this without leaving an unbonded NH group.

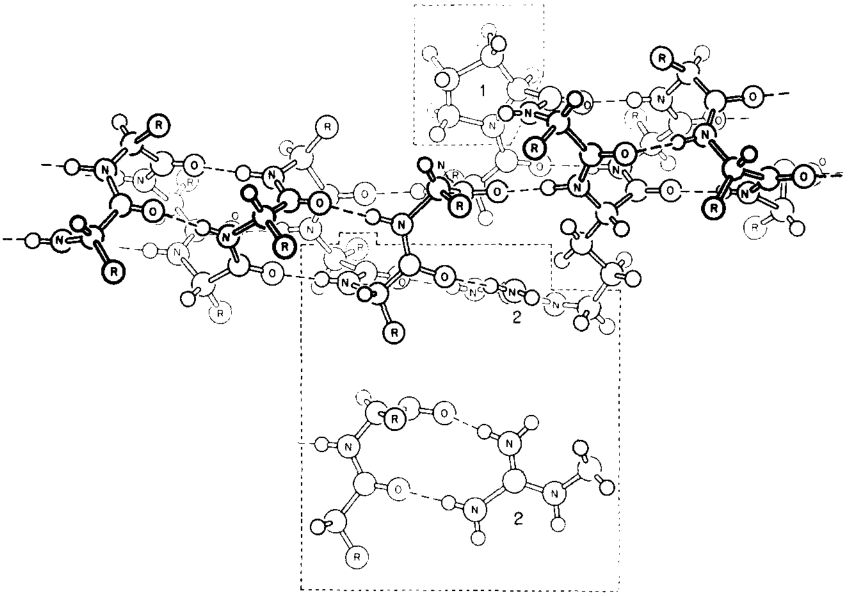

Fig. 3. This illustrates how L-proline (1) may be incorporated into an α-helix immediately following a L.H. to R.H. sense change. In this instance the change of sense is stabilised only by an arginine residue (2), and the way in which this occurs is shown very clearly in the (lower) elevation drawing. This also illustrates how the arginine is well situated to form an interchain salt link. The specific sequence involved is -pro(1)-X-arg(2)-.

4. *Possible R.H. to L.H. changes of sense*

We have seen how both the incorporation of L-proline and intrachain L-cystine into an α-helix can be most easily explained on the basis of a change of sense of the α-helix from L.H. to R.H. If this is correct then the occurrence of more than one proline residue or intrachain sulphur bond into a continuous peptide chain can only be accomplished with an α-helical structure if changes of the opposite sense, *i.e.* from R.H. to L.H., also occur.

Just as for the converse change of sense there are a number of ways in which this can be achieved. Fig. 4 illustrates one such way which makes use only of the side chains of the dicarboxylic acids and their amides. The precise sequence involved is -glu-X-X-asp-asp(NH_2)-X-X-glu(NH_2)- but many other possibilities could be invoked. No requirement in the way of a particular amino-acid which necessitates this particular sense change seems to exist however. It is conceivable that the L.H. helix is the naturally preferred type, though no reason for this can be advanced. Radial distribution curves on proteins however have been said to indicate that most proteins give better agreement with the α-helix with the β-carbon atom in position 2, *i.e.*, corresponding to a L.H. α-helix for L-amino-acids[8]. However it is difficult to see any theoretical justification for this assumption. The only obvious factor which might influence the relative stability of the L.H. and R.H. α-helices is the closeness of approach of the carbonyl oxygen and the adjacent β-carbon atoms. This is considerably closer in the case of the L.H. helix (2.75 A compared with 3.5 A for the R.H. α-helix).

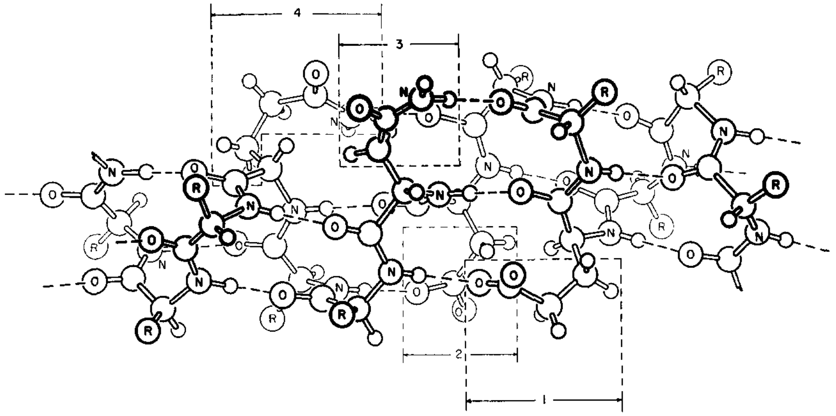

Fig. 4. One possible configuration of a R.H. to L.H. sense change of an α-helix involving the amino-acid sequence –glutamic(1)–X–X–aspartic(2)–asparagine(3)–X–X–glutamine(4)–.

DISCUSSION

At the present state of our knowledge of protein structure, it would not be profitable to discuss in great detail all the possible implications of the configurations described here. However it is worth considering a few of these inferences insofar as they can be construed as providing evidence to support the idea that the α-helix configuration is a general feature of protein structure.

From this standpoint probably the most important conclusion which can be drawn is the wide variety of possible configurational effects which can be postulated as a result of incorporating the idea of a sense change into the α-helix configuration. This provides a simple way of modifying the rigid nature of the α-helix to allow its accommodation in the more flexible type of structure which globular proteins seem to require. More specifically this idea of sense changes in the α-helix provides a simple solution to the problems associated with the incorporation of proline residues and intrachain disulphide bonds of cystine. It could also be of some help in explaining the α–β transformation which occurs when wool is stretched, since a turn of a L.H. helix and a turn of an R.H. helix if extended together would produce no resultant torque on an intermediate interchain disulphide bond.

Any change of sense of an α-helix can only be obtained at the expense of losing some main chain H-bonding. There may be circumstances where this is desirable. A number of enzyme models which have been investigated make use of a carbonyl or imino group as the specific active group. In general, however, it is probable that any carbonyl and imino groups which cannot be H-bonded to the main chain because of, *e.g.*, a sense change, will try to H-bond with any group which is available, *i.e.*,

References p. 201.

a suitable side-chain group. This additional postulate of side-chain participation in main-chain H-bonding is one that has often been invoked on chemical grounds to explain aspects of protein chemistry such as for example "masked" groups and other denaturation effects, but no specific structures demonstrating this have ever been proposed. The structures discussed in this paper show how this side-chain participation in main-chain H-bonding fits naturally with the idea of a sense change in the α-helix, and hence they provide definite specific structural possibilities for many effects which have hitherto been discussed only in general terms. A structure similar to that shown in Fig. 2 but lacking the sulphur bond could serve as a model for reversible denaturation. Breakage of the salt link by lowering the pH would disrupt the specific configuration which, however, would have some chance of being reformed if the pH were subsequently raised to a suitable value.

This type of "explanation" could be extended to cover many other aspects of the chemistry and physical chemistry of proteins but enough has perhaps been said to justify the main contention of this discussion, namely that the structure of globular proteins may well be based on the α-helix configuration and that configurational deviations from the α-helix produced by the non-equivalence of amino-acid residues need only be slight. Of these possible modifications the simplest one which can be postulated, that of a sense change, could be especially important.

ACKNOWLEDGEMENTS

This investigation was carried out during the tenure of an Arthur Amos Noyes Fellowship and the author would like to express his gratitude to many colleagues at the Gates and Crellin Laboratories of Chemistry. In particular the author would like to acknowledge many valuable critical discussions with Professor R. B. Corey and Dr. J. S. Rollett. The author is also indebted to Mrs. A. A. Oberhettinger for her preparation of the drawings on which the manuscript relies so heavily.

SUMMARY

An investigation has been made by use of atomic models of some possible simple modifications of the α-helix configuration. It is concluded that relatively slight modifications such as a change of senses of the helix permit the formulation of specific structures which circumvent some of the limitations of the simple α-helix with respect to protein structure, and furthermore that these specific structures accord well with many general ideas on protein structure which have been previously proposed on chemical and other evidence.

RÉSUMÉ

Les auteurs ont étudié, au moyen de modèles atomiques, quelques modifications possibles simples de la configuration en hélice. On peut conclure que des modifications relativement légères, telles qu'un changement de sens de l'hélice, permettent la formulation de structures spécifiques qui tournent quelques-unes des limitations de l'hélice α simple dans le domaine de la structure des protéines et en outre que les structures spécifiques sont en bon accord avec la plupart des idées générales sur la structure des protéines, qui ont été antérieurement proposées à partir de données chimiques ou autres.

References p. 201.

ZUSAMMENFASSUNG

Die Möglichkeit einiger einfachen Änderungen der α-Schraubenkonfiguration wurde an Hand von Atommodellen untersucht. Es konnte aus diesen Studien die Folgerung gezogen werden, dass relativ geringe Änderungen, wie zum Beispiel eine Abweichung in der Schraubenrichtung, die Formulierung von spezifischen Strukturen genehmigen, wodurch einige Beschränkungen der Formulierung der Proteinstruktur durch die einfache α-Schraube aufgehoben werden können. Ferner konnte festgestellt werden, dass diese spezifischen Strukturen mit vielen, auf Grund chemischer und anderer Beweise früher vorgeschlagenen, allgemeinen Theorien über die Proteinstruktur übereinstimmen.

REFERENCES

[1] L. PAULING, R. B. COREY AND H. R. BRANSON, *Proc. Nat. Acad. Sci. U.S.A.*, 37 (1951) 205.
[2] H. LINDLEY AND J. S. ROLLETT, *Biochim. Biophys. Acta*, 18 (1955) 183.
[3] F. H. C. CRICK, *Nature*, 170 (1952) 882.
[4] L. PAULING AND R. B. COREY, *Nature*, 171 (1953) 59.
[5] W. L. BRAGG, E. R. HOWELLS AND M. F. PERUTZ, *Acta Cryst.*, 5 (1952) 136.
[6] F. SANGER, L. F. SMITH AND R. KITAI, *Proc. Biochem. Soc.*, (1954) vi.
[7] H. L. YAKEL AND E. W. HUGHES, *Acta Cryst.*, 7 (1954) 291.
[8] D. P. RILEY AND U. W. ARNDT, *Proc. Roy. Soc. (London) B* 141 (1953) 93.

Received April 30th, 1955