Polymers of Tripeptides as Collagen Models

III.† Structural Relationship between Two Forms of Poly (L-prolyl-L-alanyl-glycine)

It is generally believed that the triple-helical structure of collagen is a consequence of its unusual amino-acid composition (Ramachandran, 1963; Rich & Crick, 1961) and indeed it has been found that several polytripeptides which resemble collagen in having every third amino-acid residue glycine, as well as residues of one or both of the imino acids proline and hydroxyproline, also closely resemble collagen in their X-ray diffraction patterns and probably in their molecular structures (Rogulenkova, Millionova & Andreeva, 1964; Traub & Yonath, 1966). Recently, in a communication describing X-ray studies of several polytripeptides of this type, we reported (Traub & Yonath, 1966) a collagen-like X-ray pattern for poly (L-prolyl-L-alanyl-glycine), hereafter written $(Pro.Ala.Gly)_n$. We have now made more extensive investigations of this polytripeptide and found it to have a second structural modification. This communication describes the X-ray patterns of the two forms and discusses the relationship between their structures.

Two samples of $(Pro.Ala.Gly)_n$ were used in the X-ray studies; one of average molecular weight 2000 synthesized some years ago in the Biophysics Department of the Weizmann Institute (Berger & Wolman, 1961; Wolman, 1961), and a more recent preparation, of average molecular weight 6000, kindly supplied by Professor E. R. Blout of Harvard Medical School (Bloom, Dasgupta, Patel & Blout, 1966; Oriel & Blout, 1966). Specimens of $(Pro.Ala.Gly)_n$ were photographed as powders or wet pastes on 114·6-mm and 57·3-mm diameter cylindrical powder cameras and as oriented films on a flat-plate camera. Molecular models were built from rod components (5 cm = 1 Å) produced by Cambridge Repetition Engineers, and from space-filling Courtauld components (0·8 in. = 1 Å).

Powder photographs taken of specimens of $(Pro.Ala.Gly)_n$, which had been lyophilized and kept in a desiccator over calcium chloride, show a rather diffuse pattern (Table 1), which however matches all the main features of the X-ray pattern of collagen. The fairly sharp 2·88 Å reflection corresponds to the characteristic 2·9 Å meridional spacing of the protein, the strong broad bands centred at 10 Å and 5 Å to similar regions of strong intensity concentrated mainly near the equator in collagen, and the weak reflections at 7·0 Å, 3·9 Å and 3·4 Å to the near-meridional reflections on the third and seventh layer lines in the collagen pattern.

The collagen structure is based on a non-integral screw axis which relates equivalent units by a translation of 2·9 Å and a rotation of approximately 108° (i.e. ten units in three turns) (Ramachandran, 1963; Rich & Crick, 1961). These helical parameters are derived, with the aid of helical diffraction theory (Cochran, Crick & Vand, 1952), from the facts that the meridional spacing is 2·9 Å, that it lies on the tenth layer line and that there are additional relatively strong near-meridional reflections on the third and seventh layer lines.

† Part II of this series appeared as Engel, Kurtz, Katchalski in *J. Mol. Biol.* (1966).

Observed and calculated spacings for dry form of $(Pro.Ala.Gly)_n$

I_o	d_o (Å)	hkl	d_o (Å)
s	9·2 to 11·1	100	10·31
vw	7·0	103	7·02
s	4·5 to 5·5		
vw	3·9	107	3·82
vw	3·4	117	3·38
w	2·88	0,0,10	2·88

Observed intensities (I_o) were estimated as strong (s), weak (w) and very weak (vw). Indices (hkl) were assigned to the reflections and the spacings calculated (d_o) on the basis of an hexagonal cell with $a = 11·9$ Å and $c = 28·8$ Å.

The sharp 2·88 Å reflection indicates that in $(Pro.Ala.Gly)_n$ the translation per unit of structure is practically identical with that in collagen. The $(Pro.Ala.Gly)_n$ pattern is also consistent with a 108° rotation per unit in so far as the observed 7·0 Å, 3·9 Å, 3·4 Å and 2·88 Å reflections agree quite well with values calculated for 103, 107, 117 and 0,0,10, respectively, in terms of an hexagonal unit cell with $a = 11·9$ and $c = 28·8$ Å (Table 1). However, it is by no means clear that the material is fully crystalline and, in view of the diffuseness and lack of orientation of the X-ray pattern, the determination of the a axis and the assignment of the above four reflections to layer lines are quite imprecise.

In an attempt to obtain oriented photographs of $(Pro.Ala.Gly)_n$, a stroked film was prepared from an aqueous solution of the higher molecular weight material. The resulting photographs do indeed show orientation, but of a pattern that differs appreciably from the one described in Table 1. The spacings of this new pattern and their orientations are shown in Table 2. The same set of spacings was observed in powder photographs of low molecular weight material which had been wetted. The reflections are fairly sharp, except for the broad equatorial region around 10 Å, and

Observed and calculated spacings for film of $(Pro.Ala.Gly)_n$ *grown from aqueous solution*

Orientation	I_o	d_o (Å)	hkl	d_c (Å)
Equatorial	ms	9·0 to 11·2	100	9·87
Equatorial	s	5·61	110	5·70
Equatorial	m	4·97	200	4·94
Meridional	w	4·14	102	4·20
Diagonal	m	3·60	112	3·60
Meridional	mw	3·09	003	3·10

Observed intensities (I_o) were estimated as strong (s), moderately strong (ms), medium (m), moderately weak (mw), and weak (w). Indices (hkl) were assigned to the reflections and their spacings calculated (d_o) on the basis of an hexagonal cell with $a = 11·4$ Å and $c = 9·3$ Å.

they can be indexed satisfactorily in terms of a hexagonal unit cell with $a = 11.4$ Å and $c = 9.3$ Å (Table 2).

In the light of previous studies of polypeptides related to collagen, this indexing has several structurally suggestive features. The 11.4 Å a axis is of the order of the diameters found for triple helices in several polytripeptides (Traub & Yonath, 1966), including the collagen-like form of (Pro.Ala.Gly)$_n$ described above. It therefore seems likely that each unit cell contains portions of three polypeptide chains. The c axis is close to the values found for polyproline II (9.36 Å) (Cowan & McGavin, 1955; Sasisekharan, 1959), polyglycine II (9.3 Å) (Crick & Rich, 1955) and poly (L-prolyl-glycyl-glycine) (9.3 Å) (Traub & Yonath, 1966), all of which have the same left-handed helical conformation with a translation of 3.1 Å and a rotation of 120° per amino-acid residue. The indexing scheme of Table 2 indicates that the three chains are parallel rather than twisted about each other as in collagen (Cochran et al., 1952), and the fact that 003 is the only meridional reflection may well be indicative of a threefold screw axis parallel to the c-axis. Taken together, these features therefore suggest a structure of three parallel (Pro.Ala.Gly)$_n$ chains per unit cell, each having a conformation of the polyproline II type.

We have examined possible structures of this kind with the aid of molecular models and found that, to pack satisfactorily into the unit cell, the three chains must indeed be staggered by about one-third the translation per Pro.Ala.Gly tripeptide, as would be required by a threefold screw axis.

Such close packing of chains is a strong indication of interchain hydrogen bonding. In fact, the structures proposed for collagen (Ramachandran, 1963; Rich et al., 1961) and for a number of related polypeptides (Rogulenkova et al., 1964; Traub & Yonath, 1966; Crick & Rich, 1955) include systematic hydrogen bonding of the type NH . . . O between chains. We have therefore considered the various possible ways of making such interchain hydrogen bonds in (Pro.Ala.Gly)$_n$. There are, in fact, twelve possibilities to consider, as an NH . . . O hydrogen bond may involve any of two NH groups and any of three CO groups per tripeptide (see Fig. 1), and looking from the C-terminal ends of the three chains the NH . . . O sequence may be clockwise or anticlockwise.
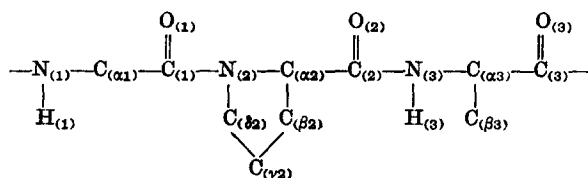

FIG. 1. Structural formula of one tripeptide unit of (Pro.Ala.Gly)$_n$ indicating notation used in text.

All six modes of hydrogen bonding involving $N_{(3)}H_{(3)}$ of the alanine residue can easily be excluded. Either the relative orientation of the CO and NH groups is incompatible with hydrogen bonding, or the C_β of alanine, or the proline ring, lies near the threefold axis and makes impossibly short contacts with atoms of other chains. By the same criteria, four of the six modes of hydrogen bonding involving $N_{(1)}H_{(1)}$ can be excluded.

Two modes of hydrogen bonding lead to reasonably satisfactory structures. These are the anticlockwise $N_{(1)}H_{(1)} \ldots O_{(1)}$ bond, which is used in the collagen I structure (Rich & Crick, 1961), and the clockwise $N_{(1)}H_{(1)} \ldots O_{(2)}$ bond, which is included in collagen II (Rich & Crick, 1961) and similar structures put forward by the Madras and King's College (London) groups (Ramachandran, Sasisekharan & Thathachari, 1962; Burge, Cowan & McGavin, 1958). The collagen-II-like structure is the more compact of the two, but without more precise criteria than we have as yet applied, one cannot exclude either possibility.

A few powder photographs were obtained which, though generally similar to the X-ray patterns described above, correspond neither to the data of Table 1 nor of Table 2. These photographs vary somewhat in their spacings. None shows either of the sharp spacings at 2·88 Å and 3·09 Å of Table 1 and Table 2, respectively, but all show a rather weak and diffuse spacing somewhere between these two values. This, and the fact that all such photographs were obtained from specimens of $(Pro.Ala.Gly)_n$ which had been exposed to water, suggest that these specimens may have partially uncoiled structures, intermediate between the collagen-like and parallel structural forms described above. However, especially in view of the lack of orientation in these X-ray patterns, their interpretation is far from certain.

Though not constituting a direct proof, our data taken together provide strong circumstantial evidence that $(Pro.Ala.Gly)_n$ in the dry state has a collagen-like triple-helical structure which uncoils on the addition of water. Such uncoiling is a necessary stage in the dissociation of a triple helix into single-stranded molecules, which is the form reported for $(Pro.Ala.Gly)_n$ in aqueous solution (Oriel & Blout, 1966). Various collagens (Josse & Harrington, 1964), as well as the collagen-like poly-tripeptides $(Gly.Pro.Hypro)_n$ (Millionova, 1964) and $(Pro.Gly.Pro)_n$ (Engel, Kurtz, Katchalski & Berger, 1966) have been found to undergo thermal transitions in aqueous solution at appropriate melting temperatures. These transitions are associated with dissociation from multi- to single-stranded conformations. It would appear that for $(Pro.Ala.Gly)_n$ the melting temperature is in fact below normal room temperatures.

It is of interest to note that two structural modifications, similar to those we have described above, have also been reported for $(Gly.Pro.Hypro)_n$ (Andreeva & Millionova, 1964).

Department of X-ray Crystallography   W. TRAUB
Weizmann Institute of Science   A. YONATH
Rehovoth, Israel

Received 14 January 1967

REFERENCES

Andreeva, N. S. & Millionova, M. I. (1964). *Soviet Physics–Crystallography (Eng. trans.)*, 8, 464.

Berger, A. & Wolman, Y. (1961). *Proc. 5th Int. Congr. Biochem.*, vol. 9, p. 82. London: Pergamon Press.

Bloom, S. M., Dasgupta, S. K., Patel, R. P. & Blout, E. R. (1966). *J. Amer. Chem. Soc.* 88, 2035.

Burge, R. E., Cowan, P. M. & McGavin, S. (1958). *Recent Advances in Gelatin and Glue Research*, ed. by G. Stainsby, p. 25. London: Pergamon Press.

Cochran, W., Crick, F. H. C. & Vand, V. (1952). *Acta Cryst.* **5**, 581.

Cowan, P. M. & McGavin, S. (1955). *Nature*, **176**, 501.

Crick, F. H. C. & Rich, A. (1955). *Nature*, **176**, 780.

Engel, J., Kurtz, J., Katchalski, E. & Berger, A. (1966). *J. Mol. Biol.* **17**, 255.

Josse, J. & Harrington, W. F. (1964) *J. Mol. Biol.* **9**, 269.

Millionova, M. I. (1964). *Biofizika*, **9**, 145. (Eng. trans. *Biophysics*, **9**, 149.)

Oriel, P. J. & Blout, E. R. (1966). *J. Amer. Chem. Soc.* **88**, 2041.

Ramachandran, G. N. (1963). *Aspects of Protein Structure*, ed. by G. N. Ramachandran, p. 39. London: Academic Press.

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

Rich, A. & Crick, F. H. C. (1961). *J. Mol. Biol.* **3**, 483.

Rogulenkova, V. N., Millionova, M. I. & Andreeva, N. S. (1964). *J. Mol. Biol.* **9**, 253.

Sasisekharan, V. (1959). *Acta Cryst.* **12**, 897.

Traub, W. & Yonath, A. (1966). *J. Mol. Biol.* **16**, 404.

Wolman, Y. (1961), Ph.D. thesis, Hebrew University, Jerusalem.