Some X-Ray Studies Concerning the Influence of Solvents on Polypeptide Structures

W. TRAUB, U. SHMUELI,† M. SUWALSKY AND A. YONATH

Department of X-ray Crystallography
The Weizmann Institute of Science, Rehovoth, Israel

X-Ray structure determinations have been made of several partially crystalline polypeptide-solvent systems. By comparing these structures with those of the dry polypeptides, an investigation has been made of the effects of solvent incorporation on the conformation and packing of the polypeptide chains.

Polyproline I forms two complexes with propionic acid. In one case there are alternating sheets of polypeptide and solvent, in the other the chains pack tetragonally and enclose columns of solvent. In both complexes there are van der Waals contacts between polypeptide chains and these have the same conformation as in the dry hexagonally packed form. In two systems involving polytripeptides related to collagen, $(Pro\text{-}Gly\text{-}Gly)_n$—formic acid and $(Pro\text{-}Gly\text{-}Pro)_n$—water, solvent incorporation does not affect either polypeptide conformation or interchain hydrogen bonding. Poly-γ-benzyl-L-glutamate in a cholesteric phase with *m*-cresol has an α-helical conformation, as in the dry form. Polylysine hydrochloride undergoes a series of reversible packing and conformational changes with increasing hydration. Its structure which is of the β-pleated sheet type at low relative humidity, changes to α-helical at 84% r.h. and to random coil in dilute solution. Polyarginine hydrochloride shows similar structural changes with variations in hydration. In this case, however, the dry form is α-helical and a transition to a β conformation occurs at water contents exceeding five molecules per arginine residue.

Apart from a few major conformational changes with fairly sharp transitions, the polypeptides studied show an invariability of conformation over a wide range of solvation and intermolecular distance.

1. Introduction

Detailed knowledge of polypeptide conformations rests almost entirely on X-ray diffraction studies of materials containing little or no solvent. Consequently, there has been uncertainty as to what extent such information can be related to the chemical and biological properties of polypeptides observed in solution.

In recent years a great variety of techniques has been developed for studying the conformations of polypeptides in solution, and a considerable body of information has been built up on this subject (Schellman & Schellman, 1964;

† Present address: Department of Chemistry, Tel-Aviv University, Israel.

Katchalski *et al.*, 1964). However, these techniques generally provide very few molecular parameters and often involve a choice between a few simple models, leaving open the possibility that more subtle conformational variations may escape detection.

The application of X-ray diffraction to the study of polypeptides in solution is in general limited by the need for some regularity of molecular arrangement. An exception is the technique of small-angle X-ray scattering, which has been used to study polypeptides in isotropic solutions (Luzzati *et al.*, 1961,1962; Saludjian & Luzzati, 1966). However, this technique involves fitting only two molecular parameters, mass per unit length and radius of gyration, to continuous scattering curves; in contrast to conventional X-ray diffraction which, depending on the number of individual spectra recorded, is capable of determining many parameters.

In fact, regular molecular packing exists to a considerable extent in various liquid-crystalline phases, as has been demonstrated by X-ray studies of a number of polypeptide–solvent systems (Robinson *et al.*, 1958; Luzzati *et al.*, 1961,1962; Saludjian *et al.*, 1963*a*,*b*). Particularly detailed and highly crystalline X-ray powder patterns were obtained by Sasisekharan (1960) from pastes of polyproline with several solvents, and we found it possible to get quite well-oriented patterns from a system of this kind (see Section 3). A powerful method for obtaining oriented X-ray patterns from fairly dilute solutions has recently been described by Parry & Elliott (1965).

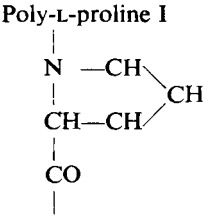
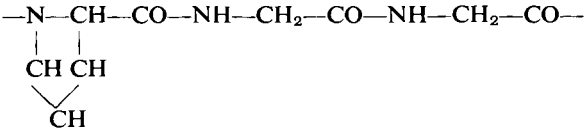
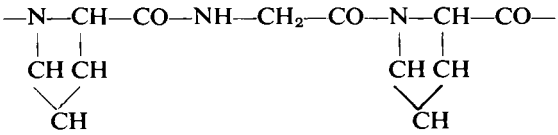
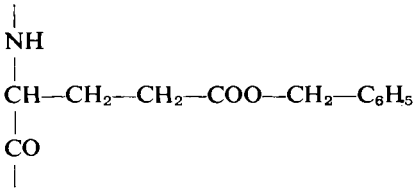
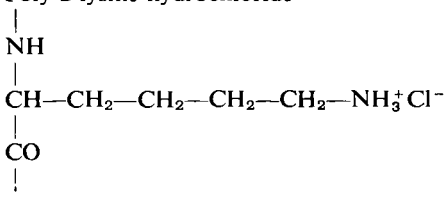
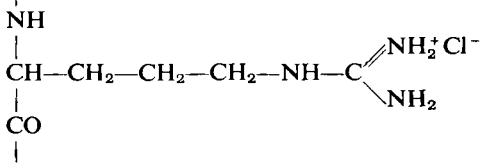
In the past few years we have determined the structures of several partially crystalline polypeptide–solvent systems and, by comparing these with the structures of the dry polypeptides, we have investigated the effects of solvent incorporation on the conformation and packing of the polypeptide chains (Traub & Shmueli, 1964; Shmueli & Traub, 1965*a*,*b*; Traub & Yonath, 1966; Suwalsky & Traub, 1966). This paper is devoted to a review and assessment of this work, much of which is here reported for the first time.

2. Materials and Methods

Table 1 lists the polypeptide–solvent systems which we investigated and indicates the molecular weight of the materials we used. Poly-L-arginine hydrochloride was supplied by Yeda, and all other polypeptides were obtained from the Biophysics Department of the Weizmann Institute, through the courtesy of Professor E. Katchalski.

Pastes, gels and oriented fibres and films of fixed solvent to polypeptide ratio were photographed in sealed glass capillaries. In some cases where the solvent was water, specimens of various degrees of hydration were prepared by suspending the material in a sealed glass capillary connected to a reservoir of saturated salt solution of known relative humidity (Shmueli & Traub,

TABLE 1. Polypeptide-solvent systems investigated

Polypeptide	Molecular weight	Solvent
Poly-L-proline I 	10,000–15,000	Propionic acid CH_3CH_2COOH
		Acetic acid CH_3COOH
		Formic acid HCOOH
Poly-(L-prolyl-glycyl-glycine) 	3,500	Formic acid HCOOH
Poly-(L-prolyl-glycyl-L-proline) 	6,000	Water H_2O
Poly-γ-benzyl-L-glutamate 	200,000	*m*-Cresol $CH_3C_6H_4OH$ (1,3)
Poly-L-lysine hydrochloride 	50,000	Water H_2O
Poly-L-arginine hydrochloride 	28,000	Water H_2O

1965*a*). Calcium chloride, magnesium chloride, sodium dichromate, sodium chloride, ammonium sulphate, potassium bromide, potassium chloride, sodium tartrate and potassium sulphate were used to obtain relative humidities of 0, 33, 52, 76, 81, 84, 86, 92 and 98% respectively.

Most powder photographs were taken with 114·6 mm and 57·3 mm diameter cylindrical powder cameras and standard X-ray units. Thin oriented specimens were photographed on a Norelco microcamera used with a Hilger microfocus X-ray tube. Thicker oriented specimens and those enclosed in capillaries were photographed on one of several flat-plate cameras with a Philips fine-focus tube. Calcite powder was used to standardize the flat-plate photographs. All photographs were taken with nickel-filtered CuK_α radiation.

Molecular models were built from rod components (5 cm = 1 Å) produced by Cambridge Repetition Engineers, and close-packing components (0·8 in = 1 Å) produced by Courtauld.

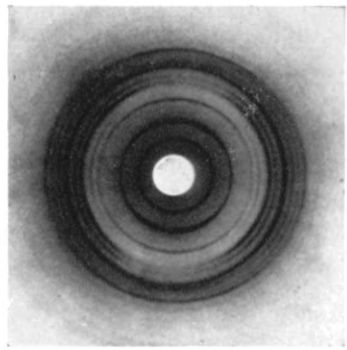
X-Ray photographs were measured with a travelling microscope as well as with a Joyce–Loebl recording microdensitometer.

3. Poly-L-proline I and Propionic Acid

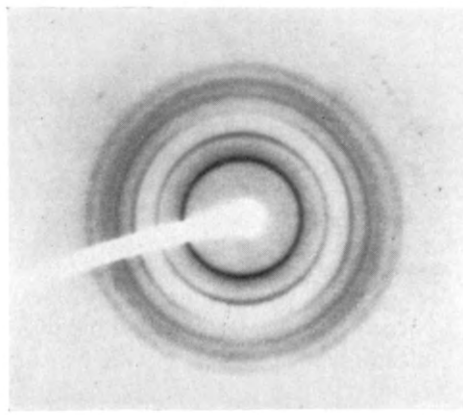
A few years ago, in the course of a structural investigation of poly-L-proline I, we drew fibres from a concentrated solution of the polypeptide in propionic acid and took a series of X-ray photographs while allowing the specimens to dry out slowly. The photographs show oriented diffraction patterns corresponding to a succession of distinct phases of high crystallinity, including a highly solvated phase (Fig. 1(a)), a less highly solvated phase (Fig. 1(b)) and dry polyproline I (Fig. 1(c)).

With the aid of such photographs, the structure of polyproline I was determined (Traub & Shmueli, 1963*a*,*b*). The polypeptide forms a right-handed helix with a translation of 1·90 Å and a rotation of 108° per proline residue. The peptide groups are in the *cis* configuration. The polypeptide chains are hexagonally close packed with 9·05 Å between adjacent chains and ten proline residues arranged in three turns along the 19·0 Å *c* axis.

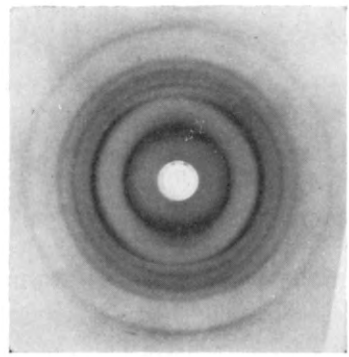
It is possible to index the X-ray patterns of the two solvated phases, in accordance with the observed orientations, in terms of unit-cell dimensions which are closely related to those of the dry form (see Tables 2 and 3). The less highly solvated form has a tetragonal unit cell with $a = 9{\cdot}13$ Å, and $c = 19{\cdot}0$ Å, and the more highly solvated, a centred orthogonal cell with $a = 9{\cdot}00$ Å, $b = 25{\cdot}1$ Å and $c = 19{\cdot}0$ Å. In both solvated forms, as in dry polyproline I fibres drawn from propionic acid, reverse orientation was observed, i.e. the *c*-axes, which are parallel to the lengths of the polypeptide chains, are perpendicular to the lengths of the fibres.



(a)



(b)



(c)

FIG. 1. X-Ray diffraction photographs of oriented fibres of (a) highly solvated polyproline I–propionic acid complex, (b) less highly solvated complex, (c) dry polyproline I. All photographs were taken with the fibre axis vertical.

TABLE 2. Observed and calculated spacings for complexes of poly-L-proline I with propionic, acetic and formic acids

	Propionic acid complex				Acetic acid complex		Formic acid complex		
hkl	d_c(Å)	Orientation	I_o	d_o(Å)	I_o	d_c(Å)	I_o	d_o(Å)	d_c(Å)
100	9·13	meridional	vs	9·16	vs	9·14	s	8·92	8·92
110	6·46	meridional	m	6·52	m	6·49	w	6·32	6·31
103	5·20	equatorial	ms	5·17	ms	5·21	m	5·17	5·16
200	4·57	meridional	m	4·57	s	4·53	m	4·47	4·46
113	4·52	diagonal	s	4·52					4·47
104	4·22	equatorial	mw	4·25	m	4·27			4·19
210	4·08	meridional	m	4·10	ms	4·06	mw	4·02	3·99
203	3·70	diagonal	m	3·74	ms	3·72	m	3·69	3·65

Observed intensities (I_o) were estimated as very strong (vs), 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_c) on the basis of tetragonal cells with $a = 9{\cdot}13$ Å, $c = 19{\cdot}0$ Å for propionic and acetic acids and $a = 8{\cdot}92$ Å, $c = 19{\cdot}0$ Å for formic acid. Note the reverse orientation.

Not only do all three forms have the same 19·0 Å repeat distance along the chain direction, but in each case there are several relatively strong reflections with index $l = 3$, indicating in accordance with helical diffraction theory (Cochran *et al.*, 1952) that 19·0 Å corresponds to three turns of the helix. As it is very unlikely that the same helical pitch and repeat distance correspond to different numbers of proline residues in the different forms, it seems clear that the two solvated phases contain polypeptide chains with the same helical conformation as was found in the dry form of polyproline I.

The three forms also have closely similar *a*-axes indicating that van der Waals contacts are maintained between adjacent molecules in the solvated phases. The small variations in the *a*-axes can be explained by the fact that the molecules are not cylindrically symmetrical, but in fact, have tenfold helical symmetry. The different packing arrangements would therefore involve some variation in the nature of the intermolecular van der Waals contacts.

Figure 2 illustrates how polypeptide chains with the same conformation would pack in the three different unit cells. In contrast to the close-packed hexagonal arrangement, the tetragonal packing provides space for columns of solvent parallel to the polyproline I helices, and the centred orthogonal cells contain alternating layers of polypeptide and solvent. Taking into account the approximately 9·0 Å diameter of the polyproline I helices, the

TABLE 3. Observed and calculated spacings for more highly solvated complex of poly-L-proline I with propionic acid

	I_o	d_o(Å)	hkl	d_c(Å)
Equator	s	12·53	020	12·55
	m	5·65	023	5·65
	m	5·05	033	5·05
	m	3·81	044	3·79
			062	3·83
First layer line	vs	8·48	110	8·47
	m	6·15	130	6·13
	ms	5·13	103	5·18
			113	5·07
	m	4·82	123	4·79
	s	4·43	133	4·40
			150	4·38
	m	4·14	114	4·14
	m	3·98	143	3·99
			124	3·98
			152	3·98
Second layer line	m	4·25	220	4·24
	ms	3·87	222	3·87
			231	3·88
	mw	3·66	203	3·67
			240	3·66
			232	3·66
			213	3·63
Third layer line	mw	2·99	310	2·98

Observed intensities (I_o) were estimated as very strong (vs), strong (s), moderately strong (ms), medium (m) and moderately weak (mw). Indices (hkl) were assigned to the reflections and their spacings calculated on the basis of an orthogonal cell with $a = 9{\cdot}00$ Å, $b = 25{\cdot}1$ Å and $c = 19{\cdot}0$ Å. Only values of hkl have been listed for which helical diffraction theory (Cochran *et al.*, 1952) indicates there could be appreciable intensity. Reflections with closely similar spacings on adjacent layer lines, particularly the first and equatorial, may sometimes be incompletely resolved. For example, the spacing observed at 4·43 Å, which has its greatest intensity on the first layer line may also have contributions from 043 ($d_c = 4{\cdot}46$ Å), 024 ($d_c = 4{\cdot}44$ Å) and possibly even 200 ($d_c = 4{\cdot}50$ Å). Other such reflections include 013 ($d_c = 6{\cdot}14$ Å), 060 ($d_c = 4{\cdot}18$ Å), 034 ($d_c = 4{\cdot}13$ Å) and 153 ($d_c = 3{\cdot}61$ Å).

observed cell dimensions allow some 3·9 Å and 3·6 Å for the widths of the columns and sheets respectively. These dimensions are consistent with monomolecular thickness of solvent and normal van der Waals contacts in the two solvated phases.

The side to side aggregation of polyproline I molecules in the highly solvated phase suggests an explanation for the observed reverse orientation. Presumably in drawing fibres the sheets, rather than individual molecules,

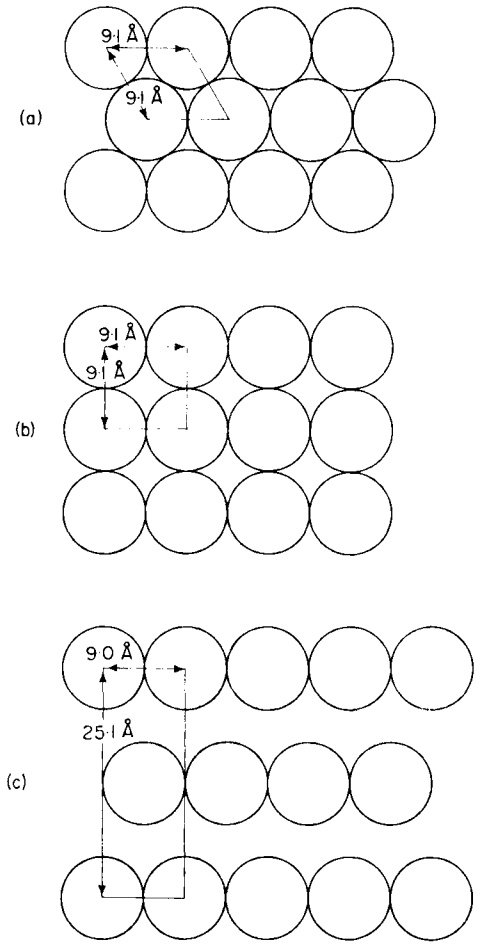


FIG. 2. Schematic illustration of packing of polypeptide chains in (a) dry polyproline I, (b) less highly solvated polyproline I–propionic acid complex, (c) more highly solvated polyproline I–propionic acid complex. All three structures are viewed along the polypeptide chains, i.e. parallel to the *c*-axis.

are oriented with their lengths parallel to the fibre axis. If the sheets are more extended perpendicular than parallel to the fairly low molecular weight polyproline I molecules, this would cause reverse orientation. In subsequent drying out there would hardly be room for molecules to rotate about axes perpendicular to their lengths. However, transitions to the tetragonal and subsequently hexagonal modes of packing could easily be brought about by rotations about, and small translations perpendicular to, the long axes of the molecules.

Sasisekharan (1960) has described crystalline X-ray powder patterns obtained from pastes of polyproline I with acetic and with formic acids. These resemble very closely the X-ray pattern of the tetragonal polyproline I —propionic acid phase described above. Indeed, as shown in Table 2, the reflections of the acetic acid complex can be satisfactorily indexed in terms of the same cell dimensions as were found in the case of propionic acid, whereas the formic acid complex has the same c axis but a slightly smaller a axis. Presumably the smaller formic acid molecules can be incorporated with slightly closer packing of the polyproline I helices.

4. Poly-(L-Pro-Gly-Gly) and Formic Acid

The structure of poly-(L-prolyl-glycyl-glycine), hereafter written (Pro-Gly-Gly)$_n$, has been determined (W. Traub, unpublished).

(Pro-Gly-Gly)$_n$ has the monoclinic space group $P2_1$ and an approximately orthogonal unit cell with $a = 12{\cdot}2$ Å, $b = 4{\cdot}9$ Å, $c = 9{\cdot}3$ Å and $\beta = 90°$. The structure of the polytripeptide is shown schematically in Fig. 3. The chains are parallel to the c axis and form left-handed helices, with each tripeptide corresponding to an axial translation of 9·3 Å and a rotation of 360°,

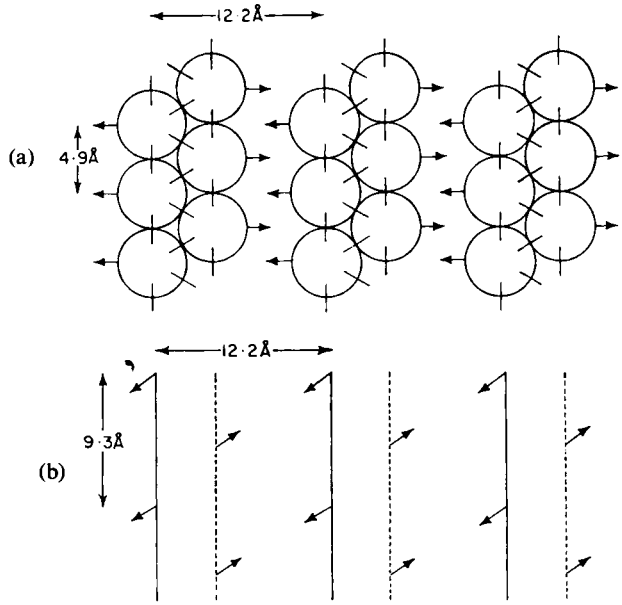


FIG. 3. Schematic illustration of the structure of (Pro-Gly-Gly)$_n$ as viewed (a) along the c-axis and (b) along the b-axis. In (a) polypeptide chains are represented by circles, pyrrolidine rings by arrows and hydrogen bonds by lines. In (b) polypeptide chains are represented by full or dashed lines and pyrrolidine rings by arrows.

approximately equally divided between the three amino acid residues. The chains thus have essentially the same conformation as has been found for polyproline II (Cowan & McGavin, 1955; Sasisekharan, 1959) and polyglycine II (Crick & Rich, 1955; Ramachandran *et al.*, 1966*b*).

The chains are held together by two NH · · · O hydrogen bonds per tripeptide to form double-layered sheets, with the two layers antiparallel. All the pyrrolidine rings of the proline residues are on the outsides of the sheets, as are the carbonyl groups—one per tripeptide—which are not involved in interchain hydrogen bonding. Each chain has a lateral separation of 4·9 Å from each of its four nearest neighbours.

X-Ray powder photographs were taken of several pastes which had been prepared from (Pro-Gly-Gly)$_n$ and various amounts of formic acid. Only one X-ray pattern of a solvated phase was observed, Table 4. This resembles the

TABLE 4. Observed and calculated spacings for (Pro-Gly-Gly)$_n$–formic acid complex

I_o	d_o(Å)	hkl	d_c(Å)	I_o	d_o(Å)	hkl	d_c(Å)
vs	13·58	100	13·47	s	4·28	011	4·34
		001	9·28			102	4·29
vw	7·50	101	7·37			111	4·08
m	6·74	200	6·73	mw	3·92	210	3·96
w	5·26	201	5·26			301	3·92
s	4·61	002	4·64	w	3·67	202	3·69
		110	4·61			211	3·58
		300	4·49	w	3·39	400	3·37
						012	3·36

Observed intensities (I_o) were estimated as very strong (vs), strong (s), medium (m), moderately weak (mw), weak (w) and very weak (vw). Indices (hkl) were assigned to the reflections and their spacings calculated (d_c) on the basis of a monoclinic unit cell, space group $P2_1$, with a = 13·5 Å, b = 4·9 Å, c = 9·3 Å and β = 94°.

pattern of dry (Pro-Gly-Gly)$_n$ in general intensity distribution, but the spacings of the reflections are all somewhat different from the corresponding spacings in the dry form.

As it did not prove possible to obtain oriented X-ray patterns of the solvated phase, the problem of indexing the reflections was approached in an indirect manner. It was assumed, from the close resemblance of the two patterns, that the first few long spacings have the same indices in dry and solvated (Pro-Gly-Gly)$_n$. This was found to imply a longer a-axis in the solvated form, but much the same c-axis. Similarly if the 4·61 Å spacing is identified with the very strong 110 at 4·55 Å in dry (Pro-Gly-Gly)$_n$, the b-axis is the same in the two forms. In fact, as is shown in Table 4, all the reflections of the solvated

phase can be indexed satisfactorily in terms of a unit cell with the same *b*- and *c*-axes as in dry (Pro-Gly-Gly)$_n$. Though the lack of orientation and limited resolution of the X-ray pattern does not allow determination of the cell dimensions with great precision, it seems clear that the *b*- and *c*-axes change little if at all.

These cell dimensions have very plausible structural implications. The 9·3 Å *c*-axis corresponds to one tripeptide unit and one turn of the polypeptide helix. The 4·9 Å *b*-axis is related to the width of the chains and to the mode of interchain hydrogen bonding. It therefore seems likely that the conformation of the chains and the hydrogen bonding between them is unaffected by the solvent. However, the sheets of polypeptide chains become slightly staggered and the distance between them increases providing space for the incorporation of formic acid, possibly in the vicinity of the unbonded carbonyl groups.

5. Poly-(L-Pro-Gly-L-Pro) and Water

Poly-(L-prolyl-glycyl-L-proline), hereafter written (Pro-Gly-Pro)$_n$, is a water-soluble polytripeptide which has been found to have an X-ray pattern closely resembling that of collagen and a very similar structure (Traub & Yonath, 1966). Three helical polypeptide chains are wound about each other to form a three-stranded coiled coil. They are held together by systematic interchain hydrogen bonds of the type NH · · · O between amino and carbonyl groups. In (Pro-Gly-Pro)$_n$ there is one such hydrogen bond per tripeptide.

It was found that, even after standing several days over phosphorus pentoxide *in vacuo*, one molecule of water per tripeptide remained bound to the polymer. We have suggested that these water molecules may form hydrogen bonds with the two carbonyl groups per tripeptide not hydrogen bonded to an amino group, and thus help to stabilize the structure by making CO · · · HOH · · · OC bridges between chains.

X-Ray photographs of (Pro-Gly-Pro)$_n$ were taken at 0, 52, 76, 86, 92 and 98% relative humidity. Those at 0 and 52% r.h. are essentially identical. They show a pattern that corresponds to a hexagonal cell with $a = 12{\cdot}5$ Å and $c = 28{\cdot}5$ Å and indicates a helical conformation with 10 tripeptide units arranged in three turns along the 28·5 Å *c*-axis (Traub & Yonath, 1966). The X-ray patterns obtained at 76% and higher relative humidities, though rather similar to that obtained at 0% r.h., are appreciably sharper and show some changes in spacing and intensity. The changes in spacing of the 100, 110 and 200 reflections in particular indicate a gradual expansion of the hexagonal unit cell with increasing humidity such that the *a*-axis increases from 12·5 Å at 52% r.h. to 13·6 Å at 98% r.h. without any appreciable change in the *c*-axis, Table 5.

TABLE 5. Observed spacings in Ångstrom units of $(Pro\text{-}Gly\text{-}Pro)_n$ axial reflections at various relative humidities

hkl	Relative humidity (%)					
	0	52	76	86	92	98
100	10·85	10·90	11·23	11·58	11·74	12·00
110	6·19		6·45	6·60	6·67	6·89
200	5·45	5·47	5·64	5·71	5·84	5·93
0, 0, 10	2·85	2·83	2·86	†	†	2·87

† Photographs of specimens at 86 and 92% r.h. were calibrated assuming the spacing of the 0,0,10 reflection to be 2·85 Å.

Photographs taken at 98% r.h. and of very wet pastes of $(Pro\text{-}Gly\text{-}Pro)_n$ show both the pattern of the hexagonal cell with $a = 13{\cdot}6$ Å and a rich pattern with many sharp lines evidently due to another phase. The latter pattern includes a strong 13·6 Å reflection and several related spacings, a possible indication that molecules, with the same conformation and diameter as in the most expanded hexagonal cell, are rearranged so as to provide additional space for water (cf. Section 3). Photographs were taken of homogenous wet pastes of $(Pro\text{-}Gly\text{-}Pro)_n$ containing 6 and 11 molecules of water per tripeptide. The specimen with 6 molecules showed the hexagonal pattern with an a-axis of 13·5 Å; the one with 11 molecules showed the pattern of the most expanded hexagonal cell as well as reflections indicating a smaller but substantial proportion of the wetter phase. From these results we estimate that the most expanded hexagonal cell contains between seven and nine molecules of water per tripeptide.

6. Poly-γ-benzyl-L-glutamate and *m*-Cresol

It was reported several years ago (Luzzati *et al.*, 1961) that a 20–35% solution of poly-γ-benzyl-L-glutamate, hereafter written PBLG, in *m*-cresol forms a cholesteric phase, in which the translation per amino acid residue appears to be about 2·0 Å as compared with 1·5 Å for the α-helix. This conclusion was based on measurement of a broad low-angle spacing, the centre of which varies with concentration, and the assumption that this corresponds to the first equatorial reflection of a fairly regular hexagonal array of molecules throughout the solution. We sought to investigate the matter further by a study of the high-angle pattern, which can generally provide more direct evidence regarding the conformation of helical structures.

Solutions of PBLG in *m*-cresol of 25 and 30% composition by weight were prepared by mixing weighed quantities of the two components and allowing

them to homogenize by standing for several weeks. Their compositions were confirmed by analysis for nitrogen. The solutions were sucked into capillaries and X-ray photographs taken with the capillaries vertical. Fairly well-oriented patterns were obtained (Fig. 4) which show broad equatorial reflections over the ranges 19–30 Å and 19·5–25 Å for the 25 and 30% solutions respectively, in agreement with the reported values. In addition, the patterns of both solutions show a near-meridional arc of medium intensity at 5·22 Å and two weak ones at 13·2 Å and 4·45 Å.

These spacings agree well with those of near-meridional streaks on the 5th, 2nd and 6th layer lines respectively of the paracrystalline dry form of PBLG (Elliott *et al.*, 1965; Parsons & Martius, 1965) which has an α-helical

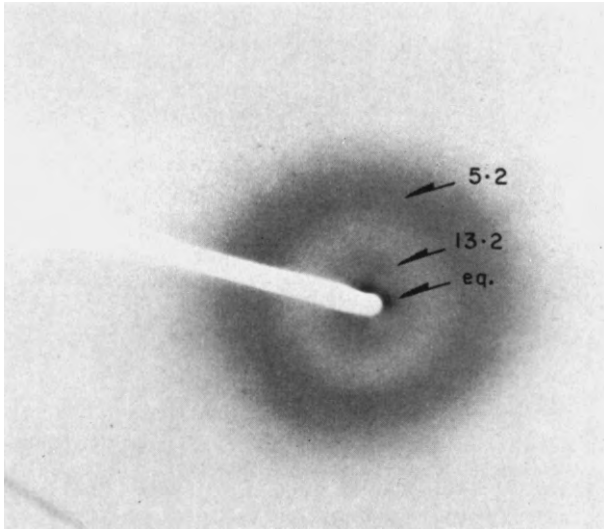


FIG. 4. Oriented X-ray diffraction photograph obtained from 25% solution of PBLG in *m*-cresol.

conformation. Evidently in the cholesteric phase the PBLG chains have this same conformation, but are packed more sparsely than had been assumed and in a rather irregular manner, as indicated by the breadth of the equatorial reflection.

Essentially the same conclusions have been reached by Parry & Elliott (1965) who investigated the PBLG–*m*-cresol cholesteric phase independently of us, and with an elegant technique were able to record the 1·5 Å meridional spacing characteristic of α-helices as well as the 5·22 Å arcs.

7. Poly-L-lysine Hydrochloride and Water

Poly-L-lysine hydrochloride is a water-soluble polypeptide which resembles basic proteins in several of its properties. Blout & Lenormant (1957) first showed from an infrared investigation that water can induce reversible configurational changes in this polypeptide. More recently, an X-ray study has been made of its molecular structure at various degrees of hydration and it has been found to undergo a remarkable series of structural changes (Shmueli & Traub, 1965*a*).

Dry polylysine hydrochloride has a pleated-sheet structure characteristic of β-polypeptides and an approximately orthogonal unit cell. Figure 5 illustrates the main features of the structure.

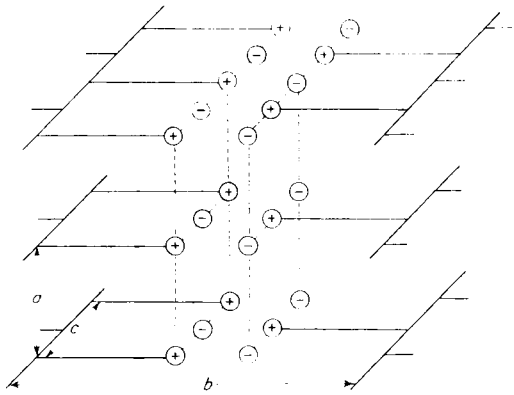


FIG. 5. Schematic illustration of structure proposed for polylysine hydrochloride at 0% relative humidity. Polypeptide chains lie with their lengths along the *c*-axis and are hydrogen bonded together to form pleated sheets parallel to *ac*-plane. The lysine side chains are roughly perpendicular to this plane and there is a regular alternation of NH_3^+ and Cl^- ions.

With increasing relative humidity over the range from 0 to 84%, the inter-sheet distance increases gradually from 15·2 to 17·0 Å, whereas the 6·66 Å repeat distance along the polypeptide chain remains unchanged. There is also a small increase from 4·62 to 4·78 Å in the 100 spacing, due either to a lengthening of the *a*-axis or increasing perpendicularity between the *a*- and *b*-axes. The changes in the X-ray pattern over this range of relative humidity are shown in Table 6.

Polylysine hydrochloride contains practically no water at 0% r.h. and about five molecules per lysine residue at 84% r.h. Water is presumably absorbed mainly around the NH_3^+ and Cl^- ions at the ends of the side chains, causing

TABLE 6. Observed spacings of polylysine hydrochloride in the β form at various relative humidities

	I_o	hkl	Relative humidity (%) 0	33	52	65	84
Equator	w	010	15·10	15·67	16·25	16·55	16·94
	vs	020	7·57	7·85	8·11	8·35	8·51
	m	030	5·08	5·24	5·42	5·54	5·65
	vw	050		3·13	3·23	3·34	
	vs	100	4·62	4·68	4·66	4·71	4·78
First layer line	vw	011		6·34	6·21	6·33	6·30
	w	021		5·16	5·13	5·23	5·25
	mw	101 111	3·78	3·78	3·77	3·80	3·82
Second layer line	mw	002	3·34	3·33	3·33	3·31	3·36

Observed intensities (I_o) were estimated as very strong (vs), medium (m), moderately weak (mw), weak (w) and very weak (vw). These varied little with relative humidity. Spacings are given in Ångström units. Features of the α form of polylysine hydrochloride appearing at 84% relative humidity have not been included in this Table. Indices (hkl) were assigned to the reflections on the basis of an orthogonal cell with, depending on the relative humidity, $a = 4{\cdot}62$–$4{\cdot}78$ Å, $b = 15{\cdot}2$–$17{\cdot}0$ Å and $c = 6{\cdot}66$ Å.

an increase in the separation between sheets with only very little alteration in the structural conformation of the sheets.

X-Ray patterns of specimens containing from five to nine molecules of water per residue no longer indicate pleated sheets, but show characteristic features of α-helices packed in hexagonal array. The interhelix distance varies with water content from 16·0 Å in a "superdried" form containing four water molecules per residue, through 16·8 Å for five molecules at 84% r.h., to 19·0 Å for nine molecules. Table 7 shows the variation in the equatorial X-ray spacings over this range of hydration.

Photographs of specimens containing 15 and 20 water molecules per residue show neither pleated-sheet nor α-helical patterns nor indeed any crystalline diffraction features. This may be due to an isotropic solution of α-helices or a change of conformation to a random-coil form which has been found to predominate in dilute solution (Applequist & Doty, 1962).

There are thus two conformational transitions, as well as packing changes in the pleated-sheet and α-helical forms, which are brought about merely by changes in the water content. The various structures are stabilized mainly by ionic forces, and the transitions can be explained in terms of increased mobility of the chloride ions with increased hydration (see Discussion).

TABLE 7. Observed spacings of polylysine hydrochloride in the α-form at various degrees of hydration

I_o	hkl	d_o	d_o	d_o	d_o	d_o
vs	100	13·90	14·50	15·28	15·91	16·40
s	110	8·01	8·39	8·67	9·12	9·47
ms	200	6·90	7·25	7·59	8·00	8·19
w	210		5·53		5·99	
No. of molecules of water/residue ..	..	4	5	6	7	9

The observed intensities (I_o) varied little with degree of hydration and were estimated as very strong (vs), strong (s), moderately strong (ms), moderately weak (mw) and very weak (vw). Observed spacings (d_o) are given in Ångström units, and the approximate water content as molecules of water per residue of lysine hydrochloride. Indices (hkl) were assigned in terms of hexagonal packing so that the interhelix distance corresponds to $2/\sqrt{3}$ of the 100 spacing.

8. Poly-L-arginine Hydrochloride and Water

Poly-L-arginine hydrochloride was first synthesized only recently (Ariely *et al.*, 1966) and its properties have as yet been studied very little. Like polylysine hydrochloride, it is a water-soluble polypeptide derived from a basic amino acid, and we were interested to investigate whether its conformation shows a similar dependence on water content (see Section 7).

X-Ray photographs of specimens containing up to about five molecules of water per arginine residue show features characteristic of α-helical structures, including a 5·4 Å layer line and a meridional 1·5 Å reflection. Increasing the water content from one half to five molecules per residue causes the *a*-axis of the hexagonal unit cell to increase from 14·4 to 15·8 Å, with no appreciable change in the 27·0 Å *c*-axis. Removal of the last half molecule of water results in a very diffuse α pattern, but on rehydration the sharp pattern reappears.

Photographs of specimens containing 5–20 molecules of water per residue show a quite different X-ray pattern, the spacings of which do not vary appreciably with hydration. The pattern includes a meridional 3·4 Å reflection, a feature commonly shown by β pleated-sheet structures, and can be indexed satisfactorily in terms of a monoclinic unit cell with $a = 9{\cdot}3$ Å, $b = 22{\cdot}1$ Å, $c = 6{\cdot}8$ Å and $\gamma = 109°$.

With the aid of molecular models, we have been able to show that these dimensions are compatible with a β pleated-sheet structure for polyarginine hydrochloride. In our model the polypeptide chains lie along the *c*-axis and the pleated sheets are parallel to the *ac*-plane (cf. Fig. 5) and staggered relative to each other along the *a*-direction. Adjacent chains in the pleated sheets are antiparallel as indicated by the 9·3 Å *a*-axis, which should be roughly half

this length in the case of parallel chains. Because of the bulky planar guanidyl groups (Ramachandran *et al.*, 1966*a*), it is not possible for the ions at the ends of the side chains from adjacent pleated sheets to overlap in the way illustrated for polylysine hydrochloride in Fig. 5. This makes the intersheet distance considerably larger than in polylysine hydrochloride, even though one would not expect the lengths of the side chains to be very different in the two polypeptides.

In our model, adjacent sheets are held together by water bridges between the ions. Furthermore, there is room for the order of 10 additional water molecules per residue to penetrate the structure, which is relatively open because of the large intersheet distance. However, the addition or removal of this water would not affect the dimensions of the structure, in accordance with our observation that the X-ray spacings do not vary over a wide range of hydration.

9. Discussion

Our studies have covered a considerable variety of polypeptide-solvent systems. They have been concerned with several solvents with different properties, structures stabilized in different ways, including van der Waals, ionic and hydrogen bonding, and in fact nearly all the conformations that have been found for polypeptides.

One general limitation of the investigation is that it has been largely confined to concentrated solutions, and it is therefore appropriate to consider evidence from other sources regarding polypeptide conformation in corresponding dilute solutions. A study of the viscosity, sedimentation, diffusion and optical rotatory properties of polyproline I in propionic acid (Steinberg *et al.*, 1960) indicates a helical conformation very similar to that found in the solid state; judging from optical rotation this is true also of solutions in acetic and formic acids. Polyproline I mutarotates to polyproline II at various rates in the three acids. Our photographs and those of Sasisekharan (1960) of the polyproline I–acid complexes do indicate a small proportion of unsolvated polyproline II, but provide no evidence of any intermediate conformation.

(Pro-Gly-Gly)$_n$ has been reported to form extended aggregates in dilute acetic acid, but to undergo little or no aggregation in concentrated formic acid (Oriel & Blout, 1966). In this case, it seems that a greater proportion of formic acid than our specimens contained might lead to a disruption of intermolecular hydrogen bonding.

(Pro-Gly-Pro)$_n$ in aqueous solution does have a multi-stranded hydrogen-bonded structure, though it is not clear whether this is identical with the the triple-helical form found in the crystalline state (Engel *et al.*, 1966).

PBLG has been reported to be α-helical in *m*-cresol solution on the basis of various physical measurements (Doty *et al.*, 1956; Yang & Doty, 1957). This

polypeptide, as well as polycarbobenzoxy-L-lysine and poly-L-glutamic acid have been studied by X-ray diffraction over a wide range of concentration in several solvents (Luzzati *et al.*, 1961,1962; Saludjian *et al.*, 1963*a*,*b*). Taking into account a revision of some of the earlier results due to improvements in low-angle X-ray scattering techniques (Saludjian & Luzzati, 1966,1967), these studies indicate α-helical conformations for all three polypeptides in both concentrated and dilute solutions. The suggestion that the ω-helical conformation exists in the quadratic phases of polycarbobenzoxy-L-lysine and poly-L-glutamic acid in dimethylformamide (Saludjian *et al.*, 1963*a*,*b*) does not seem to us adequately supported, as the reported layer-line spacing is compatible with an α-helix and our results have shown that tetragonal packing does not necessarily imply fourfold helical symmetry (see Section 3).

Perhaps the most striking fact to emerge from our studies is the invariability of conformation observed over a wide range of conditions. Even polylysine and polyarginine hydrochlorides, which show reversible α to β transitions, maintain their conformations quite precisely over considerable ranges of hydration and intermolecular distance. Polypeptide conformations do of course vary appreciably under the influence of solvents (Singer, 1962), but our studies suggest that such variations involve major conformational changes with fairly sharp transitions. However, our results are still very limited and other polypeptide–solvent systems may behave quite differently from those we have studied.

The remarkable conformational transitions shown by polyarginine and polylysine hydrochlorides evidently arise from the ionic character of these structures. A similar reversible transition has been reported for the sodium salt of polyglutamic acid (Lenormant *et al.*, 1958; Johnson, 1959). The sensitivity to hydration of these compounds can be explained in terms of repulsions, between ions fixed at the ends of the side chains, which are shielded by unbound counterions. Increased hydration would lead to increased mobility of the counterions and consequent destabilization and eventual disruption of the various structures (Shmueli & Traub, 1965*a*).

The nature of the changes in intermolecular packing caused by increased solvation in the various systems may provide information about the relative strength of polypeptide–solvent and polypeptide–polypeptide interactions. The regular increase of intermolecular distance with solvation, observed in most cases, suggests a fairly even association of polypeptide and solvent, such as might result from hydrogen bonding between specific groups. By contrast the maintenance of van der Waals contacts in the polyproline I–propionic acid system indicates that the solvent does not bind in an equivalent manner to all the proline residues and that polypeptide–polypeptide attractions are at least as strong as those between polypeptide and solvent.

$(Pro\text{-}Gly\text{-}Pro)_n$ shows both kinds of polypeptide–solvent association; an even binding of seven to nine water molecules per tripeptide and then a rearrangement which keeps the hydrated molecules in contact and provides interstices for additional water. This system seems somewhat analogous to the binding of water to collagen (Harrington & von Hippel, 1961; Berendsen & Migchelsen, 1965) and the distinction between bound and free water in crystalline proteins (Kendrew, 1962).

ACKNOWLEDGMENTS

We wish to thank Professors Ephraim Katchalski and Gerhard Schmidt for their encouragement and interest in this work, Dr Joseph Kurtz for providing us with many materials and much valuable advice, and Dr Ezra Daniel who initiated us in the study of polylysine hydrochloride. This investigation was supported by research grant GM 08608 from the National Institutes of Health, United States Public Health Service.

REFERENCES

Appelquist, J. & Doty, P. (1962). *In* "Polyamino Acids, Polypeptides and Proteins", ed. by M. A. Stahmann, p. 161. Madison: University of Wisconsin Press.

Ariely, S., Wilchek, M. & Patchornik, A. (1966). *Biopolymers*, **4**, 91.

Berendsen, H. J. C. & Migchelsen, C. (1965). *Ann. N.Y. Acad. Sci.* **125**, 365.

Blout, E. R. & Lenormant, H. (1957). *Nature, Lond.* **179**, 960.

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

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

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

Doty, P., Bradbury, J. H. & Holtzer, A. M. (1956). *J. Am. chem. Soc.* **78**, 947.

Elliott, A., Fraser, R. D. B. & MacRae, T. P. (1965). *J. molec. Biol.* **11**, 821.

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

Harrington, W. F. & von Hippel, P. H. (1961). *Adv. Protein. Chem.* **16**, 91.

Johnson, C. K. (1959). Ph.D. thesis, Massachusetts Institute of Technology.

Katchalski, E., Sela, M., Silman, H. I. & Berger, A. (1964). *In* "The Proteins", ed. by H. Neurath, Vol. 2, p. 405. New York: Academic Press.

Kendrew, J. C. (1962). *Brookhaven Symp. quant. Biol.* **15**, 216.

Lenormant, H., Baudras, A. & Blout, E. R. (1958). *J. Am. chem. Soc.* **80**, 6191.

Luzzati, V., Cesari, M., Spach, G., Masson, F. & Vincent, J. M. (1961). *J. molec. Biol.* **3**, 566.

Luzzati, V., Cesari, M., Spach, G., Masson, F. & Vincent, J. M. (1962). *In* "Polyamino Acids, Polypeptides and Proteins", ed. by M. A. Stahmann, p. 121. Madison: University of Wisconsin Press.

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

Parry, D. A. D. & Elliott, A. (1965). *Nature, Lond.* **206**, 616.

Parsons, D. F. & Martius, V. (1965). *J. molec. Biol.* **10**, 530.

Ramachandran, G. N., Mazumdar, S. K., Venkatesan, K. & Lakshminarayanan, A. V. (1966*a*). *J. molec. Biol.* **15**, 232.

Ramachandran, G. N., Sasisekharan, V. & Ramakrishnan, C. (1966*b*). *Biochim. biophys. Acta* **112**, 168.

Robinson, C., Ward, J. C. & Beevers, R. B. (1958). *Disc. Faraday Soc.* **25**, 29.

Saludjian, P., de Lozé, C. & Luzzati, V. (1963*a*). *C.r. Hebd. Séanc Acad. Sci., Paris*, **256**, 4297.

Saludjian, P., de Lozé, C. & Luzzati, V. (1963*b*). *C.r. hebd. Séanc Acad. Sci., Paris*, **256**, 4514.
Saludjian, P. & Luzzati, V. (1966). *J. molec. Biol.* **15**, 681.
Saludjian, P. & Luzzati, V. (1967). *In* "Poly-α-Amino Acids—Protein Models for Conformational Studies", ed. by G. D. Fasman. New York: Dakker (in press).
Sasisekharan, V. (1959). *Acta cryst.* **12**, 897.
Sasisekharan, V. (1960). *J. Polymer Sci.* **47**, 373.
Schellman, J. A. & Schellman, C. (1964). *In* "The Proteins", ed. by H. Neurath, Vol. 2, p. 1. New York: Academic Press.
Shmueli, U. & Traub, W. (1965*a*). *J. molec. Biol.* **12**, 205.
Shmueli, U. & Traub, W. (1965*b*). *Israel J. Chem.* **3**, 42.
Singer, S. J. (1962). *Adv. Protein Chem.* **17**, 1.
Steinberg, I. Z., Harrington, W. F., Berger, A., Sela, M. & Katchalski, E. (1960). *J. Am. chem. Soc.* **82**, 5263.
Suwalsky, M. & Traub, W. (1966). *Israel J. Chem.* **3**, 247.
Traub, W. & Shmueli, U. (1963*a*). *Nature, Lond.* **198**, 1165.
Traub, W. & Shmueli, U. (1963*b*). *In* "Aspects of Protein Structure", ed. by G. N. Ramachandran, p. 81. London: Academic Press.
Traub, W. & Shmueli, U. (1964). Proc. Sixth Int. Congr. Biochemistry, New York. Vol. 2, p. 194.
Traub, W. & Yonath, A. (1966). *J. molec. Biol.* **16**, 404.
Yang, J. T. & Doty, P. (1957). *J. Am. chem. Soc.* **79**, 761.