Table 4. *Silicides and germanides with transition metals of the 4th to 6th group of the period system*

Compound formula	M_3X		M_5X_3						MX		MX_2							
Structure type	$DA15$		$T1$		$T2$		$D8_8$		$B20$		$C54$		$C49$		$C40$		$C11$	
	S	G	S	G	S	G	S	G	S	G	S	G	S	G	S	G	S	G
Titanium	−		−		−		+	+	−		+	+	−	−	−	−	−	−
Zirconium	−	−	−	−	−	−	+	+	−	−	−	−	+	+	−	−	−	−
Hafnium											−		+		−		−	
Vanadium	+	+	+		−		+		−		−		−		+		−	
Niobium	−		+	+	+		+	+	−		−	−	−	−	+	+	−	−
Tantalum	−		+	+	+	+	+	+	−		−	−	−	−	+	+	−	−
Chromium	+	+	+	+	−		+		+	+	−		−		+		−	
Molybdenum	+	+	+	+	+		+		−		−		−				+	+
Tungsten	−		+		+		(+)		−		−		−				+	

+: structure is formed.
−: structure is not formed.
S = silicide; G = germanide.

whether the $D8_8$ phase could be stabilized at higher carbon contents. It seems clear, however, that the stability of the $T1$ binary phase with respect to the $D8_8$ phase increases with increasing group number in the periodic table. Further studies of the germanides will undoubtedly strengthen this interesting and important analogy between these two groups of intermetallic compounds.

References

CARLSON, O. N., ARMSTRONG, P. E. & WILHELM, H. A. (1956). *Trans. Amer. Soc. Met.* **48**, 843.

NOWOTNY, H., PARTHÉ, E., KIEFFER, R. & BENESOVSKY, F. (1954). *Mh. Chem.* **85**, 255.

NOWOTNY, H., SEARCY, A. W. & ORR, J. E. (1956). *J. Phys. Chem.* **60**, 677.

PARTHÉ, E. (1957). *Powder Metallurgy Bulletin.* **8**, 23.

Acta Cryst. (1958). **11**, 17

The Crystal Structure of L-Leucyl-L-Prolyl-Glycine*

BY YUEN C. LEUNG AND RICHARD E. MARSH

Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Passadena, California, U.S.A.

(*Received* 15 *January* 1957)

The crystal structure of the tripeptide leucyl-prolyl-glycine 'mono'-hydrate has been determined by Fourier and least-squares analysis of complete three-dimensional intensity data from copper radiation. The crystals are monoclinic with space group $P2_1$; the unit-cell dimensions are: $a_0 = 9{\cdot}44$, $b_0 = 6{\cdot}72$, $c_0 = 12{\cdot}10$ Å, $\beta = 100{\cdot}2°$. With the exception of a twist required by the presence of the proline ring, the peptide chain is in a highly extended configuration. A surprising feature is the presence of only approximately 80% of a water molecule of crystallization per molecule of tripeptide. In addition, one atom in the pyrrolidine ring of the proline residue is disordered, being located with apparently equal probability on either side of the plane of the other four ring atoms.

Introduction

Accurate determinations of the crystal structures of amino acids and simple peptides are of fundamental importance in arriving at the configurations of polypeptide chains in protein molecules. Such determinations yield information concerning the dimensions of the various components of the polypeptide chain, the methods of packing of the side chains, and the role of hydrogen bonding in determining the spatial arrangements of the chains. The confidence which can be placed in predictions of these various structural features of a protein—and, hence, the confidence with which the structure of a protein as a whole can be predicted—depends directly on both the number and the accuracy of the experimental results derived from the simpler compounds.

The amino acid proline is an important constituent

* Contribution No. 2149 from the Gates and Crellin Laboratories of Chemistry. The work described in this article was carried out under a contract (Nonr-220(05)) between the Office of Naval Research and the California Institute of Technology.

of many proteins; for example, approximately 12% of the amino-acid residues in gelatin and collagen are proline and another 10% of the residues are hydroxyproline. Structurally, proline and hydroxyproline are especially important. The nitrogen atoms in these residues have no attached hydrogen atoms and hence cannot participate in hydrogen bonding; furthermore, the steric effect of the five-membered pyrrolidine ring imposes definite restrictions on the relative orientations of neighboring residues. Proline and hydroxyproline thus play a unique role in determining the configurations which polypeptide chains assume in protein molecules.

Leucyl-prolyl-glycine was selected for detailed structural analysis for several reasons: as a tripeptide with proline in the central position, it should yield specific information concerning the effect of a proline residue upon the relative steric arrangement of the residues adjacent to it; it contains the leucyl residue, the structure of which has never before been determined by X-ray diffraction; and it will yield information concerning the configuration and manner of packing of the five-membered proline ring in peptides and proteins.

Experimental

Unit cell and space group

A specimen of L-leucyl-L-prolyl-glycine was obtained from Dr James R. Vaughan, Jr., of the American Cyanamid Company.

Leucyl-prolyl-glycine crystallizes in the form of small needles upon slow evaporation of cold aqueous solution. Weissenberg and precession photographs taken around the needle axis exhibit monoclinic symmetry, with the symmetry axis b oriented in the direction of the needle axis. The crystals show a pronounced cleavage along the (100) plane. The dimensions of the unit cell were obtained from single-crystal rotation photographs taken with the Straumanis technique; they were found to be

$$a_0 = 9{\cdot}442 \pm 0{\cdot}002,\quad b_0 = 6{\cdot}724 \pm 0{\cdot}001,$$
$$c_0 = 12{\cdot}105 \pm 0{\cdot}001\ \text{Å},$$
$$\beta = 100^\circ\ 11' \pm 2',$$
$$(\lambda \text{Cu}\ K\alpha_1 = 1{\cdot}54050\ \text{Å})$$

in which the uncertainties are standard deviations derived from least-squares treatment of the data.

The observed density of the crystals, measured by flotation, is 1·320 g.cm.$^{-3}$.* The density calculated on the basis of two molecules of the tripeptide and two molecules of water in each unit cell is 1·332 g.cm.$^{-3}$.

The water content of the crystals was determined by micro-analytical techniques, and was found to correspond to somewhat less than one molecule of water per molecule of the tripeptide. Several weighed crystals were dried *in vacuo* to constant weight, first at room temperature and then at elevated temperatures. At room temperature there was no loss in weight; at 56° C. the loss in weight was 4·33%, corresponding to 0·72 molecule of water per molecule of tripeptide; at 76° C. the total loss was 4·97%, corresponding to 0·83 molecule of water per molecule of tripeptide or 1·66 molecules of water per unit cell. The density of the crystals calculated on the basis of 1·66, rather than two, molecules of water is 1·319 g.cm.$^{-3}$, in nearly exact agreement with the observed value. Thus both the observed density and the analytical data indicate that under ordinary atmospheric conditions the crystals of leucyl-prolyl-glycine contain less than one molecule of water per molecule of tripeptide.

The only systematic absences observed in the X-ray data are ($0k0$) reflections with k odd; the indicated space group is thus $P2_1$.

Intensity data

Complete three-dimensional intensity data for Cu $K\alpha$ radiation were collected on multiple-film equi-inclination Weissenberg photographs taken about the a and b axes. All layer lines having inclination angles less than 45° were recorded (h from 0 to 8 and k from 0 to 6). Of the 1837 independent planes in the sphere of reflection 1697 were observed on the photographs. For photographs around the b axis, needle crystals about 0·3 mm. in diameter and 1 mm. in length were used. For photographs around the a axis, the needles were cut down to form nearly cubic specimens about 0·3 mm. on an edge. Since the crystals deteriorated after a few days exposure to X-rays, it was necessary to use a different crystal for the photography of nearly every layer line.

Visual estimations of the intensities of the reflections were made independently by both authors; in nearly all cases the two estimations agreed within 10%. The intensities were corrected for Lorentz and polarization factors, but no correction for absorption was considered necessary. The corrected intensities from all sets of photographs were correlated to bring them on to the same relative scale.

Approximate temperature and scale factors were obtained by Wilson's method (Wilson, 1942), and the corrected intensities were put on an absolute scale; the temperature factor B was found to be 3·17 Å^2. The atomic scattering factors of McWeeny (1951) were used throughout the investigation.

Derivation of the structure

Since the asymmetric unit of leucyl-prolyl-glycine contains 21 atoms (exclusive of hydrogen atoms) with similar X-ray scattering powers and since the molecule itself has little internal symmetry, no preliminary attempt was made to derive the approximate structure by means of two-dimensional analysis. Instead,

* Based on the experience with other related compounds, we estimated a probable error of about 0·007 g.cm.$^{-3}$ in the observed density.

the three-dimensional Patterson function was calculated immediately.

The Patterson function was sharpened by dividing each coefficient with

$$\sum_{i=1}^{21} f_i^2 \exp\left[-2B \sin^2\theta/\lambda^2\right]$$

and the peak at the origin was removed. In addition the coefficients were multiplied by the modification factor

$$\left(\frac{2\sin\theta}{\lambda}\right)^4 \exp\left[-\left(4{\cdot}4\frac{\sin\theta}{\lambda}\right)^2\right]$$

(Waser & Schomaker, 1953; Shoemaker, Donohue, Schomaker & Corey, 1950). The resulting three-dimensional plot possessed several encouraging and helpful features. A low vector density along the line $u = w = 0$ indicated that few, if any, pairs of atoms have the same x and z parameters; hence, the (010) projection, which in this space group is centrosymmetric, was expected to be well resolved. Another outstanding feature was the presence of three pronounced peaks about 1·4 Å from the origin, a distance characteristic of interactions between pairs of bonded atoms. Furthermore, the three vectors from the origin to the centers of these peaks are nearly co-planar. Fig. 1

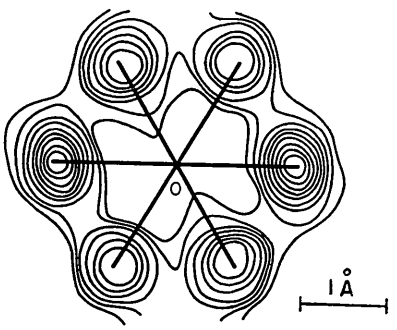
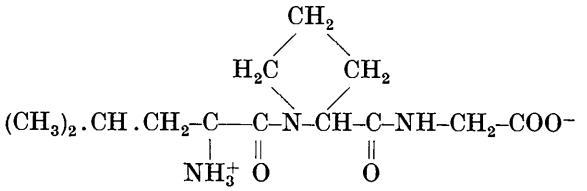


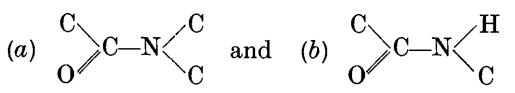
Fig. 1. The peaks around the origin of the Patterson function.

shows the plane through the Patterson function containing these three vectors and those related to them by the center of symmetry; the ends of these vectors form a nearly regular hexagon.

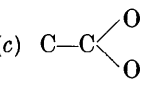
In the molecule of leucyl-prolyl-glycine



the two peptide groups



and the terminal carboxyl group



are expected to be planar. In an effort to determine the orientations of these groups, a further search was made for peaks representing second- and third-bonded neighbors. Several possible orientations for the four-membered (*c*) and five-membered (*b*) groups satisfied the Patterson function, but only one plausible orientation for a planar group of six atoms (*a*) was found; this orientation was assumed to be that of the peptide group connecting the leucyl and prolyl residues.

With this group of six atoms as a starting point, the superposition method (Shoemaker, Barieau, Donohue & Lu, 1953) was applied in the hope of finding the positions of additional atoms. However, even after sixfold superposition, there still remained a large number of peaks into which many apparently reasonable structures could be fitted. Indeed, at this stage so many structures appeared to be possible that it did not seem feasible to continue with the superposition method.

The next attack on the determination of a satisfactory trial structure was made with the aid of molecular models (Corey & Pauling, 1953*a*). Several plausible structures having satisfactory packing and hydrogen-bond arrangements were considered; some of these were eliminated on the basis of the observed cleavage along (100) or the high observed intensity of the 201 reflection. The remaining possibilities were tested in more detail by careful comparison of observed and calculated *h0l* intensities.

In formulating approximate atomic coordinates for these trial structures use was made of the sixfold overlap map. Thus, coordinates were assigned which resulted not only in satisfactory bond distances and bond angles but which were also consistent with the positions of the overlaps. As a further check on the assigned coordinates, all predicted interatomic vectors were checked against the original Patterson map and the parameters were adjusted so as to give the best fit.

Some of these trial structures looked so promising that several stages of refinement by successive computations of Fourier projections on the (010) plane were made before it was decided that the structure was unacceptable. In the course of these trials, the *R*-factor for the *h0l* data would often decrease steadily in about five stages from about 65% to about 45%. It would then be found that further small adjustments in positional parameters could not be expected to improve the agreement between observed and calculated structure factors, especially for some of the low-order reflections for which agreement remained poor. In addition, it usually became apparent that the parameters were refining toward a molecule that was structurally unacceptable.

After several unsatisfactory trials, a structure was

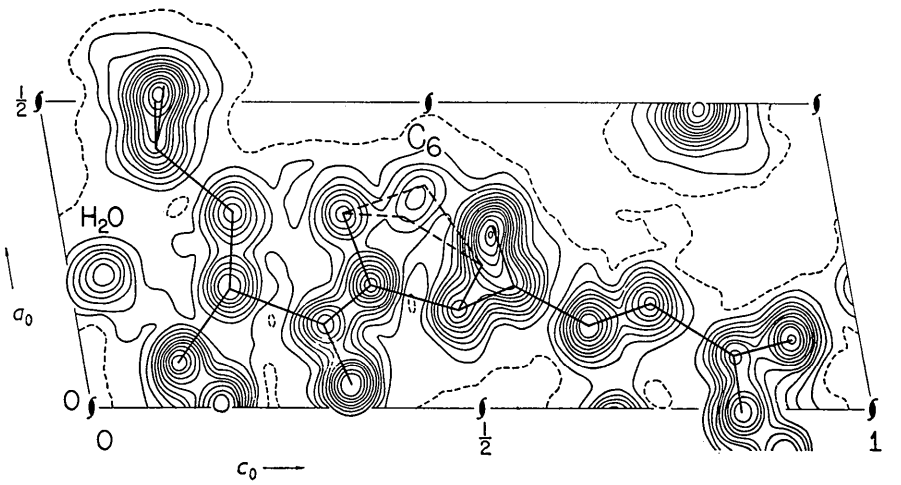


Fig. 2. The final (010) Fourier synthesis. Contours are drawn at intervals of 1 e.Å^{-3}; the dashed lines represent 1 e.Å^{-3}.

found that had none of the objectionable features of its predecessors. When refined by means of the (010) projection, the first few stages of the refinement proceeded at approximately the same rate as with the previous structures, but with the difference that subsequent stages continued to converge so as to reduce the R-factor for the $h0l$ data to a value of 28%.

The final (010) Fourier projection of electron density for this structure is shown in Fig. 2. It is well resolved and contains no spurious peaks. With two exceptions, the peaks for all the atoms have their anticipated heights. The peak for carbon atom C_6 in the proline ring is much lower than those of the other carbon atoms and is also very much elongated. This matter, however, was not taken into consideration until the refinement of the three-dimensional data. The other anomalous peak, which represents the oxygen atom of the water molecule, is only 80% of the usual height for an oxygen peak. In a difference Fourier plot computed on the assumption of a whole oxygen atom in the water molecule, there is a negative peak equivalent to about 20% of an oxygen atom. These results, together with the results of the analysis for water and the density measurement, indicate that statistically only about 80% of a water molecule is present in each asymmetric unit. The introduction at this point of only 80% of an oxygen atom reduced the R-factor for the $h0l$ data from 28% to 24%; further refinement of the x and z parameters reduced it to 21%.

The next step was the determination of the y parameters. As a start, they were calculated from the x and z parameters and the generally accepted values of the bond lengths and bond angles. In this regard, most of the ambiguities that arose from the two-dimensional nature of the (010) projection were resolved by consideration of the three-dimensional Patterson plot and the reasonable packing of the molecules. There remained, however, ambiguities in the assignment of y parameters to the two oxygen atoms of the carboxyl group and the two terminal carbon atoms of the leucyl side-chain.

Before proceeding to computations based on three-dimensional data, it seemed desirable to attempt a partial refinement of the y parameters on the basis of the $0kl$ reflections. In the projection parallel to the a axis, the molecule—and in particular the carboxyl group—was expected to be reasonably well resolved, and there were high hopes that this projection would eliminate the ambiguities concerning the y parameters. However, this did not prove to be the case. The $0kl$ structure factors calculated for both of the alternate orientations of the carboxyl group showed approximately the same quality of agreement with the observed data. In addition, Fourier and difference Fourier maps calculated on the basis of both orientations were inconclusive; even the difference map calculated on the basis of what later proved to be the wrong orientation failed to indicate that the oxygen atoms were misplaced. The conclusion was therefore reached that refinement of the non-centrosymmetric projections could not be carried much further, and attention was turned to the use of the three-dimensional data.

One pair of the two alternative orientations for the carboxyl group and for the terminal isopropyl group of the leucine residue seemed to be indicated by certain features of the Patterson plot. These orientations were tentatively selected and the selection was tested by calculation of the structure factors for planes of the type $h1l$ and $h2l$. Discrepancies between observed and calculated values for these structure factors were so large as to indicate that the assignment of the y parameters was incorrect. A three-dimensional difference Fourier map was then calculated in which the data for only 17 low-order reflections with large discrepancies were included. This map showed clearly that the assumed orientation of the carboxyl group

was wrong and, in addition, indicated that the alternative orientation, in which the carboxyl group is nearly coplanar with the neighboring peptide group, was probably correct. The indications concerning the correct configuration of the leucyl side-chain were less clear. However, the two carbon atoms involved have nearly the same x and z coordinates, so that the difference between the two possible configurations is small.

It is perhaps surprising that the (100) difference map failed to give any conclusive information concerning the arrangement of the two oxygen atoms of the carboxyl group, whereas the crude three-dimensional difference map, which, like the projection, is non-centrosymmetric, clearly indicated the correct arrangement. A few plausible explanations may be mentioned. In the first place, the reflections chosen for the three-dimensional difference map were hand-picked, only those for which the phase angles were fairly certain being used; on the other hand, all $0kl$ reflections were included in the calculation of the two-dimensional difference map. In addition, there is some overlap of atoms in the (100) projection; although this overlap does not involve the carboxyl group, it probably makes the structure factors less sensitive to changes within this group.

With the resolution of the ambiguity concerning the arrangement of the oxygen atoms of the carboxyl group, further refinement of the y parameters was carried out through the calculation of structure factors and difference maps for the (100) projection. In the course of these refinements, the correct arrangement of the carbon atoms of the leucine side-chain became apparent and the R-factor for the $0kl$ reflections was reduced from 29% to 17%. This seemed to be as far as the refinement could be carried by means of projections, and the structure appeared to be near enough to the correct one to be refined efficiently by the use of the complete three-dimensional intensity data.

Three-dimensional refinement of the parameters

The three-dimensional refinement of positional and temperature-factor parameters included three stages of least-squares calculations interspersed with the synthesis of three difference-Fourier maps. Structure-factor and least-squares calculations were carried out on the ElectroData Corporation's Datatron computer, with the program developed by Pasternak (1956); this program includes individual anisotropic temperature factors in the structure-factor routine but does not permit least-squares adjustments on them. The difference maps were synthesized on conventional IBM equipment with the M-card system devised by Prof. V. Schomaker.

The first set of structure factors was calculated from the positional parameters obtained from the (100) and (010) projections; a single isotropic temperature factor was used for all atoms, and, for computational convenience, the oxygen atom of the water molecule was put in with one-half (rather than 80%) weight. The resulting R factor was 24%. A difference map was next calculated from these structure factors. The corrections in positional parameters were obtained from the difference map with the method described by Cruickshank (1950). To these calculated shifts, an n shift (Shoemaker *et al.*, 1950) with $n = 1{\cdot}8$ was applied to correct for the incomplete convergence of a non-centrosymmetrical structure. Changes ranging up to 0·07 Å in the y parameters were indicated, while the maximum changes in the x and z parameters were about 0·02 Å. In addition, the oxygen atom of the water molecule fell on a large positive peak; the height of this peak once more indicated that approximately 80% of a water molecule is present at each site. In all subsequent calculations, the atomic form factor of carbon, rather than oxygen, was used to represent the oxygen atom of the water molecule; again, this approximation was made purely for computational convenience. Subsequent difference maps showed no significant positive or negative regions about the water molecule, other than those associated with changes in parameters.

The revised parameters obtained from the difference map were used as the basis of a structure-factor and least-squares calculation. The structure factors showed significant improvement over the previous set, the R factor dropping from 24% to 19%. The average parameter shift indicated by the least-squares calculation was about 0·03 Å, and the shifts were more or less random in direction compared to those indicated by the previous difference map. In deriving new positional parameters, the least-squares shifts were multiplied by a factor of $\frac{1}{2}$. This $\frac{1}{2}$ shift, rather than an n shift with n equal to about 1·8 (Shoemaker *et al.*, 1950) was chosen because the weighting function which was used (Hughes, 1941) did not seem to be appropriate*; at this stage of refinement it is apparent that the major cause for the disagreements between observed and calculated structure factors is not the uncertainties of observations, as is implied by the weighting scheme of Hughes, but rather the incorrectness of the model used in the calculation of structure factors.

At this stage individual anisotropic temperature-factor parameters were introduced for all 21 atoms. These parameters were derived from the first difference map by the method of Leung, Marsh & Schomaker (1957). In this method, the curvatures and heights at the atomic positions were used to evaluate the six temperature-factor parameters and one scale-factor for each atom. Thus, 21 different scale factors were derived. These ranged in value from 0·91 to 1·08;

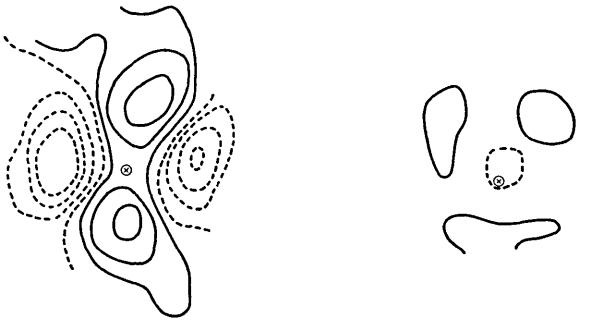
* The least-squares program is such that it was inconvenient to change the weighting function between refinement stages. Accordingly, the weighting function which was chosen for all of the refinement stages was the one we believed to be most appropriate during the final stages—that is, the function proposed by Hughes.

the average was only slightly different from unity, and no correction to the scale factor was made at this time. The temperature factor for an atom at (x, y, z) is of the form $\exp[-(\alpha h^2+\beta k^2+\gamma l^2+\delta hl+\varepsilon hk+\eta kl)]$; for the equivalent atom at $(\bar{x}, y+\frac{1}{2}, \bar{z})$ the coefficients ε and η are of opposite signs.

A second set of structure factors and a least-squares refinement were then calculated, including the individual anisotropic temperature-factor parameters in the structure factors. The R factor for these structure factors was 15·5%, and the average least-squares shift in positional parameters was about 0·02 Å. This set of calculated structure factors was also used in the preparation of a second difference map. In general, the directions of the shifts indicated in the difference map agreed with those given by the least-squares refinement; the magnitudes of the shifts were small and averages of the two results were taken, using an n shift of 1·8 for the difference map and 1·0 for the least squares. The revised positional parameters gave reasonable bond lengths and angles, except for two surprising features associated with the carbon atom C_6. First, the bond C_6–C_7 was only 1·45 Å in length; second, C_6 was nearly coplanar with the other four atoms of the proline ring, whereas the corresponding atom in hydroxyproline (Donohue & Trueblood, 1952) is 0·4 Å out of the plane of the other ring atoms.

The second difference map also indicated that the assignment of the anisotropic temperature-factor parameters had been essentially correct for all atoms except C_6. For example, a comparison of Figs. 3(*a*) and 3(*b*) shows that the positive and negative peaks around the oxygen atom O_2 had practically disappeared in only one stage of refinement. However, in the case of C_6 there was little improvement. As shown in Fig. 4(*b*), the environment of C_6 was but little changed from that shown in Fig. 4(*a*), in spite of the assignment of a relatively large anisotropic temperature factor ($\Delta B_{\max.} = 2$ Å^2). By extrapolation, it appeared that the best value of $\Delta B_{\max.}$ would be about 9 Å^2. This corresponds to a root-mean vibrational amplitude of 0·34 Å from the average position—clearly an improbable value.

All these apparent anomalies pointed to one explanation: that is, that the coordinates derived for C_6 do not correspond to the actual position of a single atom but rather that they represent the average position of two half-atoms, located with equal probability on either side of the plane containing the other four atoms of the proline ring. This implies that the

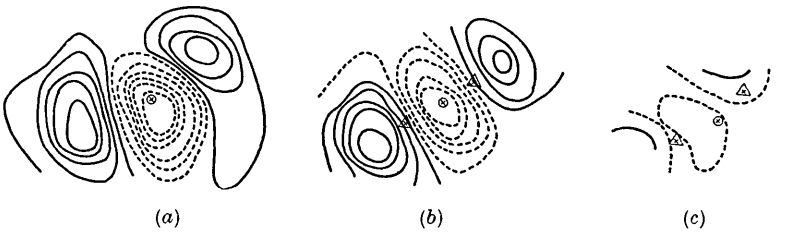


(*a*) (*b*)

Fig. 3. A portion of the difference map around O_2 at two stages of refinement: (*a*) before and (*b*) after anisotropic temperature-factor corrections. Contours are at intervals of $\frac{1}{4}$ e.Å^{-3}; the zero and negative contours are dashed.

proline ring can assume, with equal probability, either of two configurations. Each individual proline ring would thus be non-planar, and the distances from the two C_6 atoms to their neighbors would be appreciably longer than the distances observed for the average structure. Furthermore, if the magnitude of the displacement of these two half-atoms from their average position were large enough, the electron density could not be adequately represented by a single atom with large 'temperature' anisotropy. In the subsequent calculation of the structure factors, these two half carbon atoms, designated as C_6' and C_6'', were positioned symmetrically at a distance of 0·5 Å on either side of the plane of the average proline ring. Each of them was assigned tentative anisotropic temperature factors; the principal vibration of each atom was assumed to be in the direction normal to the plane defined by that atom and the two adjacent atoms, C_5 and C_7.

At this stage, the contributions of the hydrogen atoms were also calculated. Their positional parameters were obtained by geometrical considerations. The C–H, N–H, and O–H bond lengths were assumed to be 0·9 Å and the assumed bond angles corresponded to tetrahedral or plane trigonal configurations of the atoms. The disorder associated with the two con-



(*a*) (*b*) (*c*)

Fig. 4. A portion of the difference map around C_6 at three stages of refinement: (*a*) before and (*b*) after anisotropic temperature factor corrections; (*c*) final difference map, with C_6 represented by two half-atoms whose positions are indicated by triangles. Contours are as in Fig. 3.

figurations of the proline ring should require that two sets of half-hydrogen atoms be assigned to the atoms C_5, C_6, and C_7; for the purpose of computational convenience, however, only one set of full hydrogen atoms, at mean positions, was taken. For the same reason full weight was given to the hydrogen atoms of the water molecule. The calculated positions of the hydrogen atoms were checked by reference to the second difference map; in general, they fell on peaks corresponding to about $\frac{1}{2}$ e.Å^{-3}. The coordinates used for the hydrogen atoms are given in Table 1.

Table 1. *Coordinates assigned to the hydrogen atoms*

	x	y	z
$H_1(C_2)$	0·266	0·499	0·770
$H_2(C_2)$	0·155	0·631	0·709
$H_3(N_1)$	0·063	0·273	0·655
$H_4(C_4)$	0·065	0·157	0·492
$H_5(C_5)$	0·306	0·031	0·612
$H_6(C_5)$	0·204	−0·103	0·552
$H_7(C_6)$	0·430	−0·036	0·492
$H_8(C_6)$	0·339	−0·163	0·445
$H_9(C_7)$	0·393	0·139	0·351
$H_{10}(C_7)$	0·293	0·001	0·309
$H_{11}(C_9)$	0·228	0·269	0·181
$H_{12}(C_{10})$	0·370	0·505	0·294
$H_{13}(C_{10})$	0·282	0·651	0·215
$H_{14}(C_{11})$	0·375	0·523	0·070
$H_{15}(C_{12})$	0·571	0·332	0·103
$H_{16}(C_{12})$	0·547	0·305	0·219
$H_{17}(C_{12})$	0·443	0·228	0·125
$H_{18}(C_{13})$	0·591	0·686	0·116
$H_{19}(C_{13})$	0·574	0·688	0·234
$H_{20}(C_{13})$	0·477	0·807	0·151
$H_{21}(N_3)$	0·103	0·486	0·057
$H_{22}(N_3)$	0	0·378	0·112
$H_{23}(N_3)$	0·044	0·582	0·148
$H_{24}(O_5)$	0·150	−0·042	0·093
$H_{25}(O_5)$	0·194	−0·128	−0·004

A new set of structure factors and a least-squares refinement were then calculated; the hydrogen-atom contributions were included in those reflections for which $\sin\theta \leq 0{\cdot}7$. The R factor was reduced from 15·5% to 13·2%. The average shift in positional parameters, however, was still of the order of 0·02 Å. Each of the two half-atoms shifted about 0·06 Å and the oxygen atom of the water molecule shifted about 0·04 Å.

The positional parameters were adjusted using an n-shift of 1·0, and a final set of structure factors was calculated. A new scale factor was then determined; it differed by only 5% from that originally derived by Wilson's method. The resulting R factor was 12·9%. As an indication of the agreement between observed and calculated structure factors at various stages of refinement, there are listed in Table 2 the values of R obtained for all observed reflections (n in number) having h, k, or l in common; these values, together with $\Sigma|F_o|$, are listed for each such h, k, or l. A comparison of the final calculated and observed structure factors is given in Table 3. The final atomic parameters of the atoms in one asymmetric unit are listed in Tables 4 and 5; the standard errors listed in Table 4 were derived from the residuals in the last least-squares refinement.

Table 2. *The R-factors at different stages of refinement*

	R%				
h	(a)	(b)	(c)	ΣF_o	n
0	18·3	17·3	13·7	952·0	107
1	22·4	14·3	11·0	1967·2	204
2	22·0	13·8	12·3	2047·6	207
3	22·1	13·5	11·5	1613·4	195
4	24·1	14·9	12·7	1246·3	179
5	26·9	17·0	14·9	1082·5	178
6	23·9	14·9	12·8	877·5	159
7	28·5	18·0	16·4	608·6	141
8	27·8	19·6	14·8	491·7	123
9	29·6	18·6	14·8	308·0	94
10	28·4	25·9	16·7	187·9	64
11	31·4	24·2	20·6	96·5	40
12	21·0	53·8	21·3	13·6	6
k					
0	19·6	16·9	14·4	2194·3	248
1	18·8	12·8	11·8	2336·7	266
2	22·0	14·5	11·4	2194·5	269
3	26·0	15·6	12·7	1700·2	255
4	28·0	17·2	13·8	1249·9	224
5	29·9	15·7	12·2	893·8	190
6	31·5	19·2	15·7	588·9	148
7	31·2	21·7	16·9	241·3	66
8	40·5	18·1	15·7	93·2	31
l					
0	19·6	13·0	9·8	840·1	81
1	24·7	15·7	11·0	1376·8	161
2	21·1	14·6	11·8	1239·2	160
3	19·8	14·5	11·9	1286·6	156
4	22·8	14·4	10·6	1217·0	149
5	24·9	16·4	12·8	1018·9	149
6	19·9	13·7	11·3	1040·6	141
7	23·3	16·9	14·6	691·5	126
8	24·1	14·3	11·6	682·7	119
9	26·0	14·7	14·7	603·3	113
10	28·7	21·6	17·1	474·9	96
11	31·2	19·5	17·9	376·6	80
12	31·8	20·8	20·6	289·1	67
13	31·4	26·9	25·1	176·0	48
14	39·4	27·5	28·2	133·8	38
15	33·7	17·9	40·5	45·7	13
All reflections	23·6	15·6	12·9	11492·8	1697

(a) Parameters from two-dimensional refinement.
(b) Parameters after the second least-squares refinement.
(c) Final parameters.

From the final set of calculated structure factors, portions of a difference map were calculated. These portions included regions around C_6' and C_6'' and around the other atoms whose shifts in the last least-squares refinement had been relatively large. The region around C_6' and C_6'' in the final difference map is shown in Fig. 4(*c*). It clearly indicates that the separation of C_6 into two half-atoms has, indeed, resulted in a much better agreement between the calculated and observed electron densities. No significant change in positional parameters was indicated on this map; accordingly,

Table 3. *Observed and calculated structure factors for leucyl-prolyl-glycine*

The five columns in each group contain the values, reading from left to right, of h, $10|F_o|$, $10|F_c|$, $10A_c$, and $10B_c$

Table 3 (*cont.*)

Table 4. *Final positional parameters and their standard errors*

	x	y	z	σx	σy	σz
C_1	0·0863	0·4947	0·8375	0·0009	0·0016	0·0006
C_2	0·1695	0·5120	0·7421	0·0010	0·0017	0·0007
C_3	0·2016	0·3494	0·5695	0·0009	0·0017	0·0005
C_4	0·1593	0·1739	0·4917	0·0009	0·0017	0·0006
C_5	0·2339	−0·0153	0·5332	0·0011	0·0020	0·0006
C_6'	0·3649	−0·0289	0·4768	0·0019	0·0034	0·0015
C_6''	0·3102	−0·0924	0·4426	0·0021	0·0033	0·0016
C_7	0·3177	0·0737	0·3637	0·0011	0·0017	0·0006
C_8	0·1380	0·3454	0·3133	0·0009	0·0012	0·0005
C_9	0·1936	0·3829	0·2066	0·0008	0·0014	0·0005
C_{10}	0·3236	0·5253	0·2253	0·0010	0·0016	0·0007
C_{11}	0·4239	0·5220	0·1405	0·0010	0·0012	0·0004
C_{12}	0·5061	0·3284	0·1475	0·0012	0·0022	0·0009
C_{13}	0·5276	0·6945	0·1624	0·0012	0·0022	0·0009
N_1	0·1344	0·3611	0·6555	0·0008	0·0016	0·0006
N_2	0·2019	0·2052	0·3835	0·0008	0·0014	0·0005
N_3	0·0726	0·4638	0·1224	0·0007	0·0013	0·0005
O_1	0·9934	0·3585	0·8328	0·0007	0·0012	0·0005
O_2	0·1127	0·6223	0·9131	0·0008	0·0013	0·0007
O_3	0·2984	0·4650	0·5550	0·0007	0·0012	0·0005
O_4	0·0378	0·4554	0·3368	0·0006	0·0011	0·0004
O_5	0·2164	0·9785	0·0429	0·0011	0·0020	0·0009

Table 5. *Final anisotropic temperature factor parameters*

The temperature factors are in the form of

$$\exp(-B_0 \sin^2\theta/\lambda^2).\exp[-\tfrac{1}{4}(\alpha h^2a^{*2} + \beta k^2b^{*2} + \gamma l^2c^{*2} + \delta hla^*c^* + \varepsilon hka^*b^* + \eta klb^*c^*)],$$

where $B_0 = 3{\cdot}17$ Å^2.

	α (Å^2)	β (Å^2)	γ (Å^2)	δ (Å^2)	ε (Å^2)	η (Å^2)
C_1	0·248	0·273	−1·499	0·088	0·713	0·127
C_2	1·105	−0·629	−1·377	0·390	−1·615	−0·286
C_3	0·443	1·262	−2·378	0·156	0·474	0·082
C_4	0·366	0·191	−0·931	0·129	−0·121	0·222
C_5	1·467	1·845	−1·260	0·859	−0·296	−0·065
C_6'	0·065	0·399	0·016	0·064	0·321	0·160
C_6''	0·349	0·051	0·077	0·473	0·098	0·237
C_7	1·230	0·940	−1·256	1·557	0·246	0·088
C_8	0·864	−1·280	−2·205	−0·018	−0·014	−0·246
C_9	−0·192	−0·568	−1·719	0·097	0	0
C_{10}	0·430	−0·021	−0·824	0·151	−0·262	−0·046
C_{11}	1·276	0·806	−1·576	0·450	0·100	0·017
C_{12}	1·100	1·100	1·100	0·387	0	0
C_{13}	1·100	1·100	1·100	0·387	0	0
N_1	0·124	0·859	−0·549	0·044	0	−1·022
N_2	0·598	0·245	−1·978	0·345	0·077	1·328
N_3	0·247	0·248	−2·010	0·087	0	0
O_1	1·514	0·129	−0·836	1·458	−0·306	−0·573
O_2	1·322	0·173	−1·535	0·759	−0·160	−1·266
O_3	1·763	0·352	−0·628	2·186	0	0
O_4	0·713	0·097	−1·569	0·151	0·167	1·036
O_5	0·760	0·760	0·760	0·267	0	0

the refinement of parameters was deemed to be complete.

Discussion of the structure

Molecular configuration

The bond distances and angles in leucyl-prolyl-glycine calculated from the parameters listed in Table 4 are given in Table 6 and in Fig. 5. A packing drawing of one molecule is shown in Fig. 6; schematic drawings of the entire structure, viewed in the a and b directions, are shown in Fig. 7. In general, the values for

Table 6. *Intramolecular bond distances and bond angles*

C_1–C_2	1·512 Å	O_1–C_1–O_2	125·4°
C_3–C_4	1·519	O_1–C_1–C_2	118·2
C_4–C_5	1·497	O_2–C_1–C_2	116·4
C_5–C_6'	1·517	C_1–C_2–N_1	114·5
C_5–C_6''	1·507	C_2–N_1–C_3	122·2
C_6'–C_7	1·527	N_1–C_3–O_3	123·2
C_6''–C_7	1·480	N_1–C_3–C_4	115·3
C_8–C_9	1·499	O_3–C_3–C_4	121·4
C_9–C_{10}	1·542	C_3–C_4–C_5	113·3
C_{10}–C_{11}	1·515	C_3–C_4–N_2	111·2
C_{11}–C_{12}	1·510	C_4–C_5–C_6'	105·9
C_{11}–C_{13}	1·511	C_4–C_5–C_6''	107·8
		C_4–N_2–C_7	113·3
C_3–N_1	1·314	C_4–N_2–C_8	120·6
C_8–N_2	1·339	C_5–C_6'–C_7	104·2
		C_5–C_6''–C_7	107·1
C_2–N_1	1·454	C_6'–C_7–N_2	103·6
C_7–N_2	1·458	C_6''–C_7–N_2	103·2
C_4–N_2	1·452	C_7–N_2–C_8	126·1
C_9–N_3	1·492	N_2–C_4–C_5	103·7
		N_2–C_8–O_4	122·4
C_1–O_1	1·262	N_2–C_8–C_9	118·6
C_1–O_2	1·247	O_4–C_8–C_9	118·9
C_3–O_3	1·236	C_8–C_9–N_3	107·9
C_8–O_4	1·272	C_8–C_9–C_{10}	111·6
		N_3–C_9–C_{10}	111·6
		C_9–C_{10}–C_{11}	117·9
		C_{10}–C_{11}–C_{12}	110·6
		C_{10}–C_{11}–C_{13}	109·4
		C_{12}–C_{11}–C_{13}	109·9

the bond distances and angles are close to those found in similar compounds.

The standard deviations in positional parameters, as calculated from the residuals of the last least-squares refinement (Table 4), lead to calculated standard

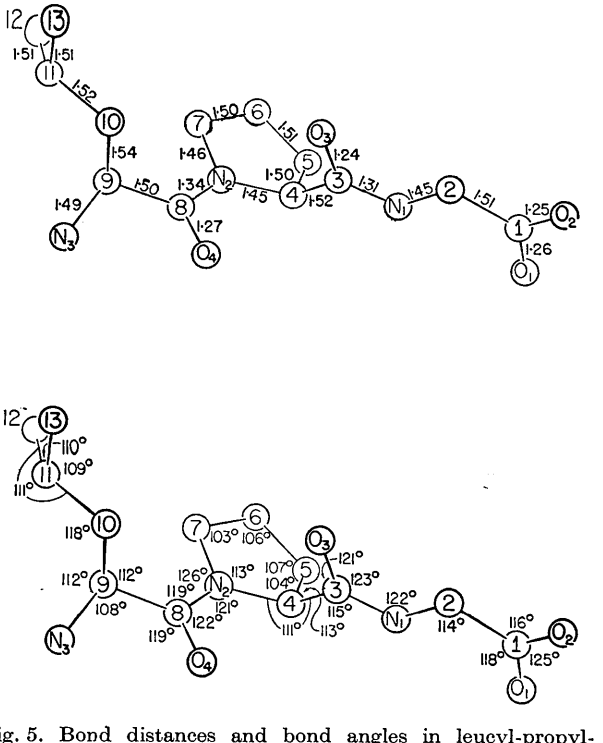


Fig. 5. Bond distances and bond angles in leucyl-propyl-glycine. Average values involving C_6' and C_6'' are indicated.

deviations, for most of the atoms, of about 0·015 Å in the bond distances and 1·0° in the bond angles, and to limits of error of about 0·03 Å and 2·0°. (For distances and angles involving atoms C_6' or C_6'', the

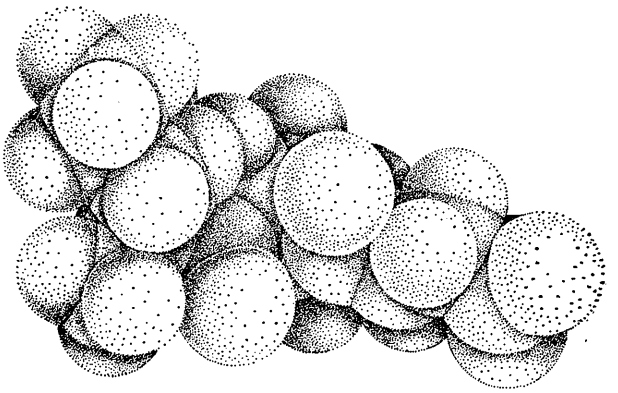


Fig. 6. A packing drawing of one molecule of leucyl-prolyl-glycine. The orientation of the molecule is the same as in Fig. 5.

calculated limits of error are approximately 0·05 Å and 3·3°; for distances and angles involving atoms C_{12}, C_{13} or O_5—the water molecule—the values are approximately 0·04 Å and 2·7°.) Subjectively, it is our opinion that the calculated limits of error are reasonable ones, and we have confidence that the bond distances and angles reported in Table 6 are accurate within these limits. Some reservation might be made for cases involving C_6' and C_6'', as the considerable overlap of electron density for these two half atoms could well increase the limit of error to, perhaps, 0·07 Å and 5°.

For purposes of discussion, the leucyl-prolyl-glycine molecule may be divided into three parts: (1) the carbon-atom side chain and the amino nitrogen atom of the leucine residue (atoms C_9, C_{10}, C_{11}, C_{12}, C_{13}, and N_3; see Figs. 6 and 7), (2) an approximately planar part consisting of the six-membered leucyl-prolyl peptide group and the proline ring (atoms C_4, C_5, C_6, C_7, C_8, C_9, N_2 and O_4), and (3) a second nearly planar part consisting of the prolyl-glycyl peptide group and the terminal carboxyl group of the glycine residue (atoms C_1, C_2, C_3, C_4, N_1, O_1, O_2, and O_2).

The dimensions of the side chain of the leucine residue differ from the expected values in two respects. In the first place, the average value for the C–C distance, not only in the leucine group but in the entire molecule, is 1·513 Å—significantly smaller than the value 1·54 Å usually assigned to a C–C single bond. Similar shortenings of the C–C bonds in hydroxyproline and other amino acids were discussed by Donohue & Trueblood (1952). In addition, there is evidence for a real alternation in the bond distances C_8–C_9 (1·499 Å), C_9–C_{10} (1·542 Å) and C_{10}–C_{11} (1·515 Å). This effect was also noted by Mathieson in the cases of methionine (1952) and norleucine (1953), and is evident in threonine (Shoemaker *et al.*, 1950). Although only in the case of threonine is the magnitude of this alternation greater than the limits of error, the effect appears to be a real one for amino-acid residues having aliphatic side chains.

A second surprising feature of the leucine side chain is the value 118° found for the bond angle C_9–C_{10}–C_{11}. The difference between this value and that of the expected tetrahedral angle, 109½°, is far greater than the experimental limit of error. In Table 7 there are listed

Table 7. *Comparison of values for the bond angel C_9–C_{10}–C_{11} in various amino acids and peptides*

Compound	Angle	Probable error	Reference
Leucyl-prolyl-glycine	117·9°	0·7°	Present study
Glycyl-tyrosine.HCl	113°	—	Smits & Wiebenga (1953)
Histidine.HCl	113°	—	Donohue *et al.*, (1956)
Threonine	112·5°	0·3°	Shoemaker *et al.* (1950)
Norleucine	114°	5°	Mathieson (1953)
α-Methionine	111°	5°	Mathieson (1952)
β-Methionine	113°	5°	Mathieson (1952)
Glycyl-tryptophan	115°	0·7°	Pasternak (1956)

values for the corresponding angles found in other peptides and amino acids. Although all of these values are larger than the tetrahedral angle, none is as large as has been found in leucyl-prolyl-glycine. A plausible explanation for the widening of this angle in the present case is the steric effect of an oxygen atom which is hydrogen-bonded to the terminal nitrogen atom.

The six atoms in the leucyl-prolyl peptide group, C_4, C_7, C_8, C_9, N_2 and O_4, are coplanar within an average deviation of 0·02 Å and a maximum deviation of 0·04 Å. The dimensions of this peptide group are close to those found in other peptides (Corey & Pauling, 1953*b*) with the exception of the bond angle N_2–C_8–C_9, for which the observed value, 118·6°, is nearly 5° larger than expected. This effect may well be due to repulsion between the hydrogen atoms attached to C_9

Table 8. *Comparison of the corresponding bond distances and angles in the pyrrolidine ring in leucyl-prolyl-glycine and in hydroxyproline*

Bonds	Leucyl-propyl-glycine	Hydroxyproline
C_4–C_5	1·497 Å	1·532 Å
C_5–C_6	1·512*	1·503
C_6–C_7	1·504*	1·524
C_7–N_2	1·458	1·482
N_2–C_4	1·452	1·503
Angles		
N_2–C_4–C_5	103·7°	104·5°
C_4–C_5–C_6	106·9*	107·6
C_5–C_6–C_7	105·7*	103·9
C_6–C_7–N_2	103·4*	105·5
C_7–N_2–C_4	113·3	109·4

* The average of two values involving C_6' and C_6''.

and C_{10} and those attached to the adjacent carbon atom, C_7, in the proline ring.

The geometry of the pyrrolidine ring of the proline residue is similar to that found in hydroxyproline (Donohue & Trueblood, 1952). The four atoms C_4, C_5, C_7 and N_2 are coplanar within a maximum deviation of 0·06 Å; the fifth atom, C_6, is about 0·37 Å out of this plane. (It actually consists of two half-atoms: C_6', which is 0·44 Å out of the plane in one direction, and C_6'', which lies 0·29 Å on the other side of the plane.) In hydroxyproline, the corresponding atom is about 0·4 Å from the plane of the other atoms. The results of the present investigation seem to establish that the puckering of the proline ring occurs mainly at C_6 rather than at other positions in the ring.

The dihedral angle between the best plane of the leucyl-prolylpeptide group and that of the proline ring is approximately 7°.

A comparison of the bond distances and angles in the pyrrolidine rings of leucyl-prolyl-glycine and of hydroxyproline is given in Table 8. The major differences are in the two C–N distances and the C–N–C bond angle. These differences are due to the different natures of the two nitrogen atoms; in leucyl-prolyl-glycine, this atom, as part of a peptide group, is tertiary, whereas in hydroxyproline, which is a zwitter-

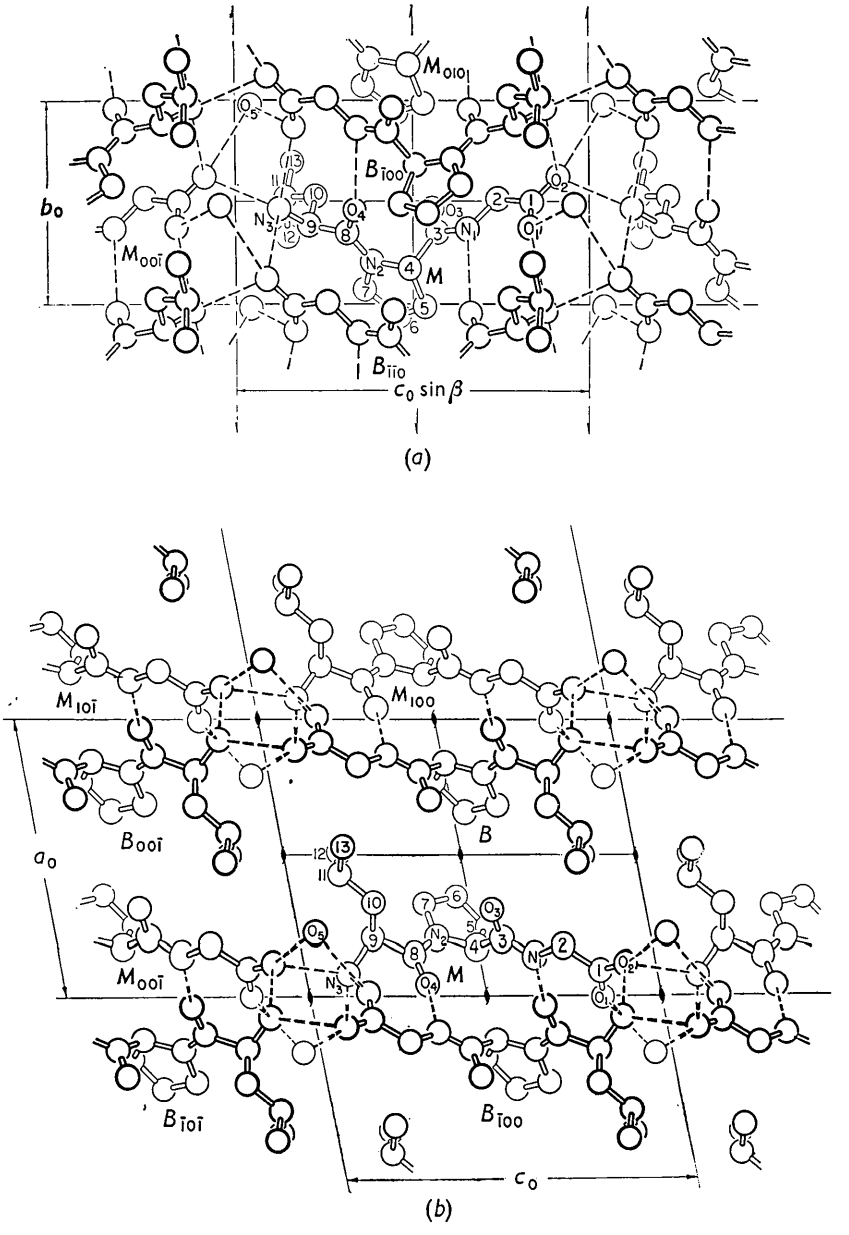


Fig. 7. Projections of the structure viewed along (*a*) the *a* axis and (*b*) the *b* axis.

ion, the nitrogen atom is quaternary. Thus, the average C–N distance found in hydroxyproline (as well as the C_9–N_3 distance in leucyl-prolyl-glycine)—1·492 Å—is typical of αC–N distances found in amino acids and in the N-terminal groups of peptides; on the other hand, the C_7–N_2 and C_4–N_2 (as well as the C_2–N_1) distances in leucyl-prolyl-glycine—approximately 1·455 Å—are close to the accepted value 1·47 Å for the αC–N distance in peptide groups (Corey & Pauling, 1953*b*). In a similar manner, the value of 109° for the C–N–C angle in hydroxyproline reflects the tendency of this nitrogen atom to be tetrahedral; the corresponding angle C_4–N_2–C_7 in leucyl-prolyl-glycine—113°—is close to the value 114° selected by Corey & Pauling (1953*b*) for the αC–N–H angle in a peptide group.

One other point of comparison between hydroxyproline and leucyl-prolyl-glycine might be mentioned. Donohue & Trueblood (1952) point out an apparent alternation in the C–C bond lengths of hydroxyproline, whereas there is little indication of such alternation involving corresponding atoms in leucyl-prolyl-glycine. Indeed, if any alternation can be inferred from the observed values for the C_3–C_4, C_4–C_5, C_5–C_6, and C_6–C_7 distances, it is in a sense opposite to that in hydroxyproline.

The dimensions of the second peptide group—that between the prolyl and glycine residues—are very close to those selected by Corey & Pauling (1953*b*); in addition, the atoms of this group (C_2, C_3, C_4, N_1 and O_3) are coplanar within an average deviation of 0·02 Å and a maximum deviation of 0·03 Å. Atoms C_1, C_2, O_1, and O_2 of the terminal carboxyl group are also coplanar within 0·01 Å; this best plane makes an angle of 1·6° with that of the adjacent peptide group.

Some remarks might be made concerning the general configuration of the leucyl-prolyl-glycine molecule. The most striking feature is the tendency of the molecule to be, as far as possible, in an extended configuration; thus, the nine main-chain atoms of the prolyl and glycine residues (C_1–C_4, N_1, N_2, and O_1–O_3) are approximately coplanar, and the amino nitrogen of the leucyl residue is close to the plane of the leucyl-prolyl peptide group. Each of the αC–N bonds is approximately *cis* to the C–O bond of the same residue. Only at the proline residue does the polypeptide chain undergo a twist; this twist occurs primarily at the C_4–N_2 bond and is required by the geometry of the pyrrolidine ring. The angle of twist—that is, the dihedral angle between the planes of the two peptide groups—is 102°, slightly larger than the value of about 90° chosen by Cowan & McGavin (1955) for their proposed structure of poly-L-proline.

There are two aspects of the extended configuration of the polypeptide chain which can be discussed. The first aspect involves the amount of twist about the C–αC bonds and, hence, the degree of coplanarity of a nitrogen atom with the neighboring peptide group. It has been pointed out by Dunitz & Robertson (1952) and by Pasternak (1956) that in all amino acids and peptides which have been investigated so far there is a strong tendency for the amino or peptide nitrogen atoms to be as close as possible to the carboxyl or carbonyl oxygen atoms of the same residue, and hence to be coplanar with the adjacent carboxyl or peptide groups. In leucyl-prolyl-glycine, N_1 is, within experimental uncertainties, exactly in the plane of the carboxyl group, and hence is as close as possible to O_1 (the O_1–N_1 distance is 2·49 Å). The terminal nitrogen atom, N_3, is 0·6 Å from the plane of the leucyl-prolyl peptide group; however, steric effects of the proline ring and the leucine side chain preclude a strict coplanarity. N_2, the nitrogen atom of the proline residue, is 0·4 Å from the plane of the prolyl-glycine peptide group; this non-planarity is presumably due to van der Waals and electrostatic repulsion between the oxygen atoms O_3 and O_4 (the O_3–O_4 distance is 3·27 Å). It is interesting to note that in hydroxyproline, in which there is no such repulsion, the nitrogen atom is only 0·05 Å from the plane of the carboxyl group.

A second aspect of the extended configuration of the polypeptide chain involves the amount of twist about the αC–N bonds. In leucyl-prolyl-glycine, a twist of about 120° from the fully-extended configuration occurs at the C_4–N_2 bond; as was pointed out previously, this twist is required by the geometry of the proline ring. On the other hand, there is no significant twist at the C_2–N_1 bond. Of the other peptides which have been investigated, twists of about 90° occur in glycyltryptophan (Pasternak, 1956), glycyl-asparagine (Pasternak, Katz & Corey, 1954), and glycyl-tyrosine (Smits & Wiebenga, 1953), whereas only very small twists occur in N,N′-diglycylcystine (Yakel & Hughes, 1952), β-glycyl-glycine (Hughes & Moore, 1949), and α-glycyl-glycine (Hughes, Biswas & Wilson, to be published). In glycyl-tryptophan and glycyl-tyrosine, the effect of this twist is to move the bulky side groups away from the polar groups of the peptide chain, leaving the latter free to form hydrogen bonds; in glycyl-asparagine, which has a polar side-chain, the situation is undoubtedly more complicated.

It appears clear that, for small peptides such as have so far been investigated, there is a strong tendency for the twist about the C–C and C–N bonds to be small and, hence, for the peptide chain to be in the fully-extended configuration.

The environment of the molecule

In Fig. 6, views of the structure along the *a* and *b* axes are shown. In these drawings the molecule labelled *M* has the coordinates given in Tables 4 and 5. Molecule *B* is related to *M* by the operation of the twofold screw axis; molecules equivalent to these in adjacent unit cells have the additional designation of a lattice translation vector. A network of hydrogen bonds in the *b* and *c* directions holds the molecules together to form sheets oriented parallel to the (100)

Table 9. *Hydrogen-bond distances and angles*

From atom x in molecule M	to atom y	in molecule	Distances N · · · O or O · · · O (Å)	Angles O · · · N–C or O · · · O · · · O (°)
O_4	N_1	$B_{\bar{1}00}$	3·185	128·9 ($O_4 \cdots N_1$–C_2) 107·0 ($O_4 \cdots N_1$–C_3)
N_3	O_2	$M_{00\bar{1}}$	2·835	122·7 ($O_2 \cdots N_3$–C_9) 109·1 ($O_2 \cdots N_3 \cdots O_2$)
N_3	O_2	$B_{\bar{1}\bar{1}0}$	2·873	100·6 ($O_2 \cdots N_3$–C_9) 112·1 ($O_2 \cdots N_3 \cdots O_1$)
N_3	O_1	$B_{\bar{1}00}$	2·830	111·1 ($O_1 \cdots N_3$–C_9) 83·5 ($O_1 \cdots N_3 \cdots O_2$)
O_5	O_1	$B_{\bar{1}00}$	2·809	81·5
O_5	O_2	$M_{00\bar{1}}$	2·936	—

plane; each of these sheets is composed of a double layer of molecules held together by hydrogen bonds in the a direction. The forces holding adjacent sheets together involve only van der Waals contacts between, principally, the side chains of the leucine residues and the pyrrolidine rings of the proline residues. This sheet structure readily explains the pronounced cleavage along (100).

All of the hydrogen atoms attached to the nitrogen atoms, as well as those of the water molecule, are involved in hydrogen bonds. The hydrogen-bond distances and angles are listed in Table 9. With the exception of the N_1–O_4 distance, all of the values are close to those observed in other peptides and amino acids (Corey & Pauling, 1953*b*); the value 3·185 Å found for the N_1–O_4 distance is significantly greater. A consideration of packing models of the structure indicates that it is impossible to shorten this distance without causing serious deformation of the structure.

The terminal nitrogen atom N_3 of the leucine residue forms three hydrogen bonds, arranged nearly tetrahedrally with respect to the N_3–C_9 bond. This tetrahedral configuration about the nitrogen atom and the near equivalence of the two C–O bonds of the terminal carboxyl group give clear indication that the molecule is present as the zwitterion. The hydrogen bonds around N_3 are all directed towards oxygen atoms of the carboxyl group; these bonds alone are sufficient to hold the structure together. Thus, although both hydrogen atoms of the water molecule are involved in hydrogen bonds, these bonds are not of major importance to the coherence of the structure. Rather, the function of the water molecules appears to be that of space filling, and it is not surprising that the rest of the structure is stable even though only about 80% of a water of crystallization is present.

The oxygen atoms O_1 and O_2 participate in two and three hydrogen bonds, respectively. The difference in environment of these two oxygen atoms is not manifested in any significant difference in either the C_1–O bond distances or the C_2–C_1–O bond angles; indeed, if the slight difference in the distances would indicate that the C_1–O_2 bond has slightly greater double-bond character than the C_1–O_1 bond, the difference in bond angles would indicate quite the opposite. It has been suggested (Pasternak *et al.*, 1954; Levy & Corey, 1941; Hughes & Moore, 1949) that dissymmetry in a CO_2^- group may occur when the two oxygen atoms are involved in different numbers of hydrogen bonds; however, it is possible that the apparent dissymmetry in these and other cases may be due, in part, to the effect on the apparent C–O bond distances of the thermal vibrations of the oxygen atoms. Thus, the oxygen atom involved in the fewer number of hydrogen bonds would presumably have the greater vibration about the C–O bond, and hence would appear to be closer to the carbon atom.

In leucyl-prolyl-glycine, the values of the temperature-factor anisotropies for atoms C_1, O_1, and O_2 lead to estimated shortenings of the C_1–O_1 and C_1–O_2 bonds of about 0·008 and 0·006 Å, respectively.* Both of these values are smaller than the standard deviation of 0·015 Å in the bond distances.

The oxygen atom of the proline residue, O_3, is not available to form hydrogen bonds; instead, it is packed snugly among three proline rings.

A list of all packing distances shorter than 4·0 Å is given in Table 10. None of these distances is surprisingly short, and on the whole the packing is rather loose. Further evidence of the looseness of the packing is the ability of atom C_6 to assume with apparently equal probability either of two positions on opposite sides of the proline ring.

We should like to express our thanks to Prof. Robert B. Corey for his continuing advice and en-

* The 'estimated shortening' of a C–O bond was arrived at by taking the maximum B value for the oxygen atom in a direction perpendicular to the C–O bond, subtracting the corresponding B value for the attached carbon atom, and calculating the r.m.s. amplitude of vibration according to the formula $\overline{u}^2 = B/8\pi^2$. The r.m.s. amplitude was then added vectorially to the observed interatomic distance to obtain the corrected distance. This method is equivalent to that described by Cruickshank (1956) when the amplitude of vibration is small; i.e., when, in Cruickshank's notation, s is much smaller than q.

Table 10. *Packing distances shorter than* 4 Å

From atom x in molecule M	to atom y	in molecule	Distance
N_3	C_1	$M_{00\bar{1}}$	3·457 Å
C_1	C_9	$B_{\bar{1}00}$	3·686
C_1	C_8	$B_{\bar{1}00}$	3·466
C_2	C_{12}	B	3·774
C_2	C_{13}	$B_{0\bar{1}0}$	3·592
C_2	C_8	$B_{\bar{1}00}$	3·635
C_2	N_2	$B_{\bar{1}00}$	3·796
N_1	N_2	$B_{\bar{1}00}$	3·890
N_1	C_4	$B_{\bar{1}00}$	3·680
N_1	C_{13}	$B_{0\bar{1}0}$	3·716
C_3	C_4	$B_{\bar{1}00}$	3·559
O_3	C_6'	M_{010}	3·617
O_3	C_6''	M_{010}	3·282
O_3	C_6'	B	3·270
O_3	C_6''	B	3·711
O_3	C_7	B	3·658
C_{13}	C_3	B	3·901
O_5	C_9	M_{010}	3·449
O_5	C_{12}	M_{010}	3·660
O_5	C_{13}	M	3·588
O_5	C_{12}	$B_{00\bar{1}}$	3·930
C_{13}	C_{12}	$B_{00\bar{1}}$	3·816

couragement during the course of this investigation. We should also like to thank Dr Robert Nathan and Dr Raphael Pasternak for their help in the use of the Datatron Computer, and the ElectroData Corporation for making the Computer available to us.

We are indebted to Dr James R. Vaughan, Jr., and the American Cyanamid Company for providing us with the sample of leucyl-prolyl-glycine.

References

COREY, R. B. & PAULING, L. (1953*a*). *Rev. Sci. Instrum.* **24**, 621.

COREY, R. B. & PAULING, L. (1953*b*). *Proc. Roy. Soc.* B, **141**, 10.

COWAN, P. M. & MCGAVIN, S. (1955). *Nature, Lond.* **176**, 501.

CRUICKSHANK, D. W. J. (1950). *Acta Cryst.* **3**, 10.

CRUICKSHANK, D. W. J. (1956). *Acta Cryst.* **9**, 757.

DONOHUE, J., LAVINE, L. R. & ROLLETT, J. S. (1956). *Acta Cryst.* **9**, 655.

DONOHUE, J. & TRUEBLOOD, K. N. (1952). *Acta Cryst.* **5**, 419.

DUNITZ, J. D. & ROBERTSON, J. H. (1952). *Rep. Progr. Chem.* **49**, 351.

HUGHES, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.

HUGHES, E. W. & MOORE, W. J. (1949). *J. Amer. Chem. Soc.* **71**, 2618.

HUGHES, E. W., BISWAS, A. B. & WILSON, J. N. To be published.

LEUNG, Y. C., MARSH, R. E. & SCHOMAKER, V. (1957). *Acta Cryst.* **10**, 650.

LEVY, H. A. & COREY, R. B. (1941). *J. Amer. Chem. Soc.* **63**, 2095.

MATHIESON, A. MCL. (1952). *Acta Cryst.* **5**, 332.

MATHIESON, A. MCL. (1953). *Acta Cryst.* **6**, 399.

MCWEENY, R. (1951). *Acta Cryst.* **4**, 513.

PASTERNAK, R. A., KATZ, L. & COREY, R. B. (1954). *Acta Cryst.* **7**, 225.

PASTERNAK, R. A. (1956). *Acta Cryst.* **9**, 341.

SHOEMAKER, D. P., BARIEAU, R. E., DONOHUE, J. & LU, C.-S. (1953). *Acta Cryst.* **6**, 241.

SHOEMAKER, D. P., DONOHUE, J., SCHOMAKER, V. & COREY, R. B. (1950). *J. Amer. Chem. Soc.* **72**, 2328.

SMITS, D. W. & WIEBENGA, E. H. (1953). *Acta Cryst.* **6**, 531.

WASER, J. & SCHOMAKER, V. (1953). *Rev. Mod. Phys.* **25**, 671.

WILSON, A. J. C. (1942). *Nature, Lond.* **150**, 151.

YAKEL, H. L. & HUGHES, E. W. (1952). *Acta Cryst.* **5**, 847.

Acta Cryst. (1958). **11**, 31

The Crystal Structure of the Hydrazine Salt of 5-Aminotetrazole

BY JOHN H. BRYDEN

Chemistry Division, Research Department, U.S. Naval Ordnance Test Station, China Lake, California, U.S.A.

(*Received* 26 *April* 1957 *and in revised form* 10 *June* 1957)

Crystals of the hydrazine salt of 5-aminotetrazole are orthorhombic, crystallizing in the space group D_2^3–$P2_12_12$. The unit-cell dimensions are $a_0 = 13{\cdot}54$, $b_0 = 9{\cdot}67$, $c_0 = 3{\cdot}86$ Å, and there are four molecules per unit cell. The structure was refined by Fourier methods, using the $hk0$ and $hk1$ data. The 5-aminotetrazole molecule is planar within the accuracy of the analysis. The hydrazine molecule exists as the NH_2–NH_3^+ ion, as shown by the hydrogen-bonding arrangement in the crystal.

Introduction

5-Aminotetrazole is a valuable intermediate in the preparation of tetrazole compounds because of its varied reactions and its ease of preparation. It is obtained by the reaction of nitrous acid with aminoguanidine (Thiele, 1892; Hantzsch & Vagt, 1901) or by the reaction of hydrazoic acid with dicyandiamide (Stollé, 1929). The first reaction forms guanylazide, which rearranges to 5-aminotetrazole: