The Crystal Structure of α-Glycylglycine*

By A. B. Biswas†, Edward W. Hughes, Brahama D. Sharma‡ and J. Norton Wilson§

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

The crystal structure of the α crystal form of glycylglycine, $^-CO_2CH_2NHCOCH_2NH_3^+$, has been investigated. The cell dimensions are $a=7\cdot70$, $b=9\cdot57$, $c=9\cdot48$ Å, $\beta=124°35'$. The space group is $P2_1/a$ with four molecules per cell. Mo $K\alpha$ rays were used. The trial structure was derived by use of part-cell Patterson methods and modified Banerjee equations. Final refinement was by block-diagonal anisotropic least-squares adjustment to an R of 12·3 and yielded standard deviations of about 0·007 Å in bond lengths. The intramolecular chemical bond lengths and interbond angles agree well with those found for the same molecule in the previously reported β crystal form but the configuration of the molecule is somewhat different. There is an angle of about 22·4° between the plane of the amide group and that of the carboxyl group. In the β crystal these groups were coplanar within experimental error. Since the α form is probably the stable form, this distortion presumably permits better molecular packing with stronger hydrogen bonds and van der Waals interactions.

In 1948 Professor Paul Ewald, as Editor of *Acta Crystallographica*, personally reviewed the paper (Hughes, 1949) discussing the modification of Banerjee's equation for sign determination. His questions and comments had a valuable clarifying effect and encouraged the trial application of the method, as described below. It is appropriate that the present long-delayed paper should appear in this issue.

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. The crystal structure of the simplest peptide, α-glycylglycine, was started in these Laboratories, among other compounds related to protein molecules, in the late 1940's. Only recently has it been possible to refine the structure to a degree suitable for publication.

When Bernal (1931) examined glycylglycine crystallographically he discovered three modifications of the substance crystallizing from n-propyl alcohol–water mixtures. From the cell dimensions which he measured, it appeared probable that the form he designated β would have the easiest structure for X-ray determination and it was accordingly chosen for our first investigation. Although the structure was indeed easily established (Hughes & Moore, 1949), the subsequent refinement of the parameters was not pushed to extreme limits because of the unfavorable cell shape. It then seemed that the α form would be more suitable for

careful study and that its structure could be arrived at without serious difficulty by using the information regarding the molecular configuration obtained from the β form, provided, of course, that the molecules had substantially the same configuration in both crystals, a question which is of considerable interest in itself.

Experimental

The crystals used were either from solutions in distilled water or from n-propyl alcohol–water mixtures. Numerous well-exposed oscillation photographs were made about the three principal crystallographic axes with use of small crystals and filtered Cu $K\alpha$ rays for determining the unit cell. The actual dimensions were taken from high-order reflections on the equators of these pictures.

The cell found (Cu $K\alpha$, $\lambda=1\cdot542$ Å) is monoclinic with $a=7\cdot70\pm0\cdot02$, $b=9\cdot57\pm0\cdot03$, $c=9\cdot48\pm0\cdot03$ Å and $\beta=124°35'\pm20'$ and contains four molecules of $C_4H_8N_2O_3$. The cell dimensions agree with those of Bernal (1931) within his rather large limits of error. The calculated density is 1·516 g.cm^{-3}. Systematic extinctions on these and later photographs, $h0l$ with h odd and $0k0$ with k odd, confirmed Bernal's space group assignment, $P2_1/a$.

For intensity measurements, Mo $K\alpha$ rays filtered through zirconium foil were used in connection with rather large crystals ground into cylinders approximately one millimeter in diameter. By using Mo $K\alpha$ radiation the quantity of data was considerably increased and the labor expended in collecting and treating these extra observations is believed to be justified by the extra resolution obtained. The multiple film technique (deLange, Robertson & Woodward, 1939) was modified by using three films interleaved with 0·001-inch copper foil, as suggested by Hughes (1941). The intensity ratio obtained for successive films was almost 4:1 for Mo $K\alpha$ radiation. Equi-inclination

---

Weissenberg pictures were prepared about **a** for the zero layer and for all odd layers up through the ninth. Another set of Weissenberg photographs was made about **b** for the zero, first and second layers. Some oscillation photographs were also prepared with Mo *Kα* radiation in a camera with 5 cm radius for correlating intensities in different layers and for observing some reflections blocked off by the beam stop of the Weissenberg camera. The intensities were estimated by visual comparison with a known intensity strip prepared by giving a series of graded exposures to one strong reflection from the same crystal. For crystals of this size with this radiation the absorption correction is negligible.

More than two thousand reflections were estimated, some of which were out as far as $\sin \theta/\lambda = 1\cdot0$; several hundred observable reflections are beyond the range of Cu *Kα* radiation. These intensities were corrected for the Lorentz–polarization factor, using the method of Lu (1943). The resulting F_{obs}^2 values were brought to an approximately absolute scale by the averaging method of Wilson (1942). More exact scale and temperature factors have been obtained by comparing the observed F's with the final calculated absolute values.

Determination of the structure

Bernal (1931) reported the crystals to be optically negative with the small refractive index for light vibrating perpendicular to (100). The orders of (100) are exceptionally intense and fall off in intensity monotonically with the scattering angle. The reflection 200 is surrounded by a region of intense diffuse blackening. It was thus clear that the crystal has an approximate layer structure parallel to (100). On the other hand, the detailed behavior of the orders of (100) showed that the structure is not exactly layered. This approximate layering suggested the use of the part-cell method (Schomaker, 1941; Booth, 1945), which requires the collection of only the F_{hkl}'s with h odd or zero and therefore reduces by nearly half the labor of collecting data.

The trial structure was arrived at by two different routes which were pursued quite independently at first.

The first of these involved the use of half-cell Patterson functions. Actually these functions were calculated when only data with $h=0$ and 1 were available. The use of all data with $h=0$ or h odd, as prescribed for the usual half-cell projection, will give a sharp cut between the two halves of the cell. If one limits the calculation to $h=0$ and $h=1$, one gets a projection of $\frac{1}{2}P(x,y,z)[1+\cos(2\pi x-\alpha)]$ where α can be selected at will. To obtain the form of the calculation one need only carry out the integration corresponding to the projection. Thus for $\alpha=0$, we have

$$P'(yz) = \frac{1}{2}\int_0^1 P(xyz)\,[1 + \cos 2\pi x]\,dx.$$

Upon substitution of the series for P one obtains

$$P'(yz) = \frac{1}{2A} \sum_k \sum_l (F_{0kl}^2 + F_{1kl}^2) \cos 2\pi(ky+lz) .$$

This is a projection into which $P(0,y,z)$ comes with unit weight while $P(\frac{1}{2},y,z)$ is eliminated; P's at intermediate x values are weighted by $(1+\cos 2\pi x)/2$. It is a part-cell projection in which the cut-off is not sharp between what is projected and what is not. For many purposes this would be of no value, but for a structure which is rather highly layered we expect almost all the Patterson peaks to be near $x=0$ or $x=\frac{1}{2}$, and these are separated by this simple treatment. For further details see works on 'generalized projections' (Cochran & Dyer, 1952).

Two such Patterson projections were prepared with $\alpha=0$ and π, respectively. It was possible to interpret these Pattersons directly, using as a molecular model the configuration found in β-glycylglycine. In this model the molecule is planar except for the terminal nitrogen atom. To fit the Patterson it was necessary to keep this plane very nearly parallel to (100), as anticipated. The long axis of the molecule had to be parallel roughly to **c** and was at approximately $y=\frac{1}{4}$. This treatment of course gives no information about the values for the x parameters although it does imply that most of them are about equal to each other.

The second route to a trial structure involved the use of the modified Banerjee equation (Hughes, 1949) in connection with the $h0l$ intensities. In the original discussion of this equation it was stated that for a centrosymmetrical structure more than $(N+1)$ observations with two indices constant would be required, where N is the number of atoms in the cell, and it was pointed out that such an extended range of data is not usually available, unless N is so small that the structure is almost obvious by inspection. It occurred to one of us (EWH) that actually N is not necessarily the number of atoms but the number of *appreciably different atom parameters* in the direction of the varying index. If two or more atoms have the same or almost the same parameter, N is accordingly smaller than anticipated from the cell content. In the present case we already expected a layer structure parallel to (100) and so it was at once suspected that many values of x parameters would be almost the same and that for varying h in the Banerjee equation N might be, effectively, quite small. The F_{h0l}'s were accordingly reduced to unitary structure factors U_{h0l}'s and an effort made to fit the equation. For this zone and space group it can be shown that the effective number of molecules per cell is only two; h is always even and the projection of one half the cell is identical with that of the other half. There are thus eighteen atoms (other than hydrogen atoms) and we should normally expect nine Banerjee constants and require at least nineteen observations, U_{h0l} with l constant. This is obviously impossible with a short *a*. However, it was found by trial that good fits could be found with only two constants and five observations with l constant, thus indicating that there are only two non-

equivalent atoms or groups of atoms with significantly different x parameters, and no atom group at $x=0$ or $x=\frac{1}{4}$. The trials were made by assuming first zero, then one, then zero and one, and finally two parameters; for each case a number of Banerjee equations were set up, always including the U's with $l=0$. By eliminating the constants from these few equations relations were formed between some U's, and signs were determined to make the relationship hold. Upon attempting to extend the system to other U's by calculating the Banerjee constants, failures (contradictory behavior) resulted for the first three trials, but with two constants and five U's per equation it was possible to extend the system throughout the entire zone and to balance the Banerjee equations within experimental error. At several stages, as more signs were added, the constants were redetermined by least squares. It should be noted that the Banerjee relationship, being linear, can determine no signs absolutely (unless $U_{0,0,0}$ is involved); when once a fit is obtained, an equally good one is had by reversing all the signs. Accordingly, one could only determine the relative signs for all U's having the same value of l. The constants obtained were $A_1 = -1.28$ and $A_2 = 0.788$. These, in turn, yield values for the parameters directly: $x_1 = \pm 0.023$ and $x_2 = \pm 0.150$ in fractions of a (or these plus $\frac{1}{4}$; with only $h0l$ intensities the two alternatives cannot be distinguished). It may be noted at this point for comparison that the parameter of atom N(1) in the final structure, referred to the same origin, is 0.1512, while the average for the remaining eight atoms is 0.0215 (the spread being from 0.001 to 0.051). The agreement is excellent; this direct process located N(1) (terminal nitrogen) to within 0.009 Å, and the average position of the remaining atoms to within 0.012 Å, all in the x direction. All but three of the relative signs turned out subsequently to be

correct, and for the wrong signs the three U's involved were small. Some efforts were made to apply the Banerjee equation with varying l but this attack was abandoned when it became clear that $N_{\text{effective}}$ was certainly greater than six. From the final structure one concludes that $N_{\text{effective}}$ is about twelve for varying l in this projection, but with two of the twelve groups at centers, so that only five constants would be required in the Banerjee equations.

Since one of the groups has x very small, its contribution to U_{h0l} varies but slowly with h (l being kept constant). The rapid variation of U_{h0l} with h is due to the second group. By inspection it was found that the maximum short range variation for any l was about ± 0.15. This represents the maximum fraction of the atoms in the second group. One accordingly concludes that since the asymmetric unit consists of nine atoms this second group in all probability is a single atom. The expression for U_{h0l} may therefore be written approximately as

$$U_{h0l} = A_l + \alpha \cos 2\pi(hx + lz)$$

where A_l is the term due to the first group of eight atoms and α is the fraction of the cell electrons in the single atom; x and z are the coordinates of this single atom, x being already fixed at about 0.150. A_l is not exactly independent of h but varies only slowly with it. This equation makes it possible to fix z, for lz appears as a phase angle in the short range variation of U_{h0l}. This phase angle must vary linearly with l, and after a few trials it appeared that the only consistent interpretation required that $z \simeq \frac{1}{6} = 0.167$. The final refined value of this parameter is 0.177, so that this direct approach fixed z for the single atom to within 0.095 Å.

The position of one atom at $x \doteq 0.150$, $z = 0.167$ having been established, the sign determination for the
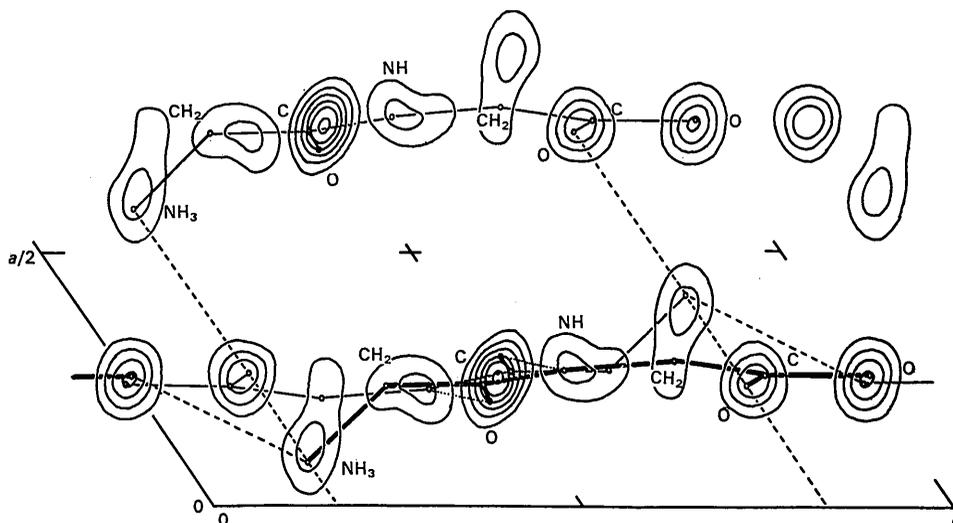
Fig. 1. Sharpened electron density projection onto (010). The layers are seen edgewise. In the bottom half the two overlapping molecules are outlined, that nearest the reader being in heavy lines. In the upper half only the farthest of the two molecules is outlined. Dotted lines indicate hydrogen bonds. Arbitrary contour interval.

entire zone follows almost automatically. The contribution of the single atom can be calculated as a function of h and l simply by drawing lines on a reciprocal lattice diagram to show the positions of plus and minus maximum contributions. One then follows along a line of constant l noting the variation of U_{h0l} in relation to the single atom contribution. If U_{h0l} increases where the latter is positive, A_l must be positive and U_{h0l} also positive. But if U_{h0l} increases where the single atom contribution is negative, A_l and U_{h0l} are negative. When once this relationship was recognized all the signs of the zone were determined in about a half hour. Three errors were made where in haste it was not noticed that A_l had itself changed in sign in the midst of a series of small U's of constant l. Fig. 1 shows the electron density projection prepared for this zone with use of the signs determined by this method. A later projection prepared with signs calculated from the final parameters does not appear significantly different and is not reproduced. There were only the six sign errors already noted, all for small U's.

At this point it appeared that it would be entirely feasible to apply these same methods to the $hk0$ data, combine the results of the two zones, and achieve a direct solution of the structure. But at this time the trial structure derived as described above from the a axis part-cell Patterson synthesis became available; it was decided not to reduce the $hk0$ data to U's but to combine the two results at hand to obtain a trial structure at once. The model used in interpreting the Patterson function was projected onto Fig. 1, using the z parameters from the Patterson interpretation and adjusting the x parameter to fit the Fourier. In making these adjustments the carboxyl and peptide groups were kept planar separately. The positions of the atoms are indicated in Fig. 1. In combining the two results it was found that the value of the x parameter from the Ba-

nerjee equations was with reference to an origin in the $h0l$ projection on a twofold screw axis. It was necessary to subtract 0·250 and change signs in order to convert the origin to a center of symmetry. The set of x, y, z coordinates obtained in this way served as a basis of further refinement.

Refinement of parameters

The first improvements were made by successive $0kl$ electron density projections. The resolution was surprisingly good. Two atoms 0·9 Å apart in the projection are resolved; half way between them $\varrho(yz)$ is only about one third of the value at the atom centers. The final projection, calculated from the signs from the final structure factor calculations, is shown in Fig. 2. The only atoms not resolved are the C(3) and O(4) of the carbonyl group in the peptide link.

Resort was then made to part-cell Fourier syntheses for resolving the carbonyl group. The structure factors were calculated by the method of Donohue & Schomaker (1949), using atomic scattering factors from *Internationale Tabellen* (1935). The part-cell syntheses, using the calculated signs, were calculated by punched-card methods. At the same time a completely diagonalized-matrix least-squares (Hughes, 1941) refinement was carried out, based on the same structure factor calculations. A separate least-squares adjustment of the scale and temperature factors was made after each step of the refinement. After the second step the hydrogen atoms were introduced at the expected positions for a zwitterion. These positions were subsequently maintained with no further refinement but were always used to allow for the scattering due to the hydrogen atoms. The hydrogen coordinates are recorded in Table 1. The temperature factor used in these early calculations was of the over-all anisotropic type (Hughes,
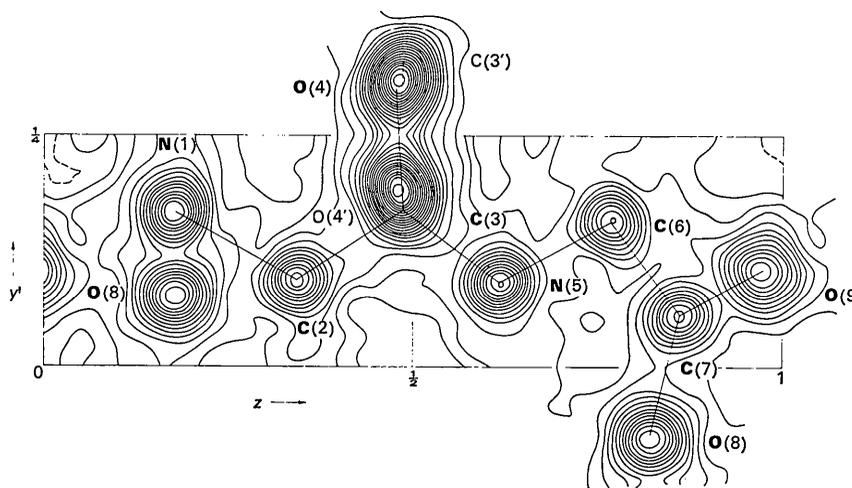
Fig. 2. Electron density projection onto (100). Because of the β-angle the two carbonyl groups, which appear almost exactly superimposed in this drawing, would be well separated in a projection normal to the planes of the molecules. Arbitrary contour interval.

1941). The final thermal parameter values found were: $A = 1.25$ Å², parallel to (100), and $A + B = 2.13$ Å², perpendicular to (100).

Table 1. *Assigned coordinates of hydrogen atoms*

Hydrogen atom no.	Bonded to atom	x	y	z
10	N(1)	0·117	0·143	0·083
11	N(1)	−0·053	0·153	0·138
12	N(1)	0·133	0·271	0·192
13	C(2)	0·182	−0·012	0·330
14	C(2)	0·396	0·074	0·372
15	N(5)	0·297	−0·013	0·617
16	C(6)	0·183	0·239	0·717
17	C(6)	0·459	0·198	0·851

On the third structure-factor calculations there were no changes in the sign for any structure factor of significant size. The second half-cell projection is shown in Fig. 3. The peaks corresponding to the heavy atoms are unusually high and sharp, as was expected; carbon atoms reach about 12 e.Å⁻², nitrogen atoms about 14 e.Å⁻² and oxygens about 15·5 e.Å⁻².

At this same stage of refinement the shifts indicated by the least-squares calculations were of the order of magnitude of the probable errors. The absolute discrepancies between the coordinates from the part-cell projection and the least-squares adjustment averaged to less than 0·001 of a cell edge. The largest discrepancy was 0·004 of a cell edge.

In the course of the refinement it was noticed that many reflections with $h \neq 0$ and k and l each equal to either 1, 2, or 3 were calculated much lower than the observed values. In all about fifteen moderately sized F's were involved. The fact that all these observed values were high, suggested a systematic error of some

kind. The b-axis Weissenberg photographs with $k = 0$, 1 and 2 were made to check these results. The intensities of these doubtful spots on the $h0l$, $h1l$ and $h2l$ pictures are much lower in relation to other reflections than they are on pictures taken about **a**. An absolutely certain reason for the discrepancy cannot be given, but it is probably a manifestation of the Renninger or 'Umweganregung' effect (Renninger, 1937).

The probability of encountering this type of double reflection increases with the density of the points in reciprocal space; with Mo $K\alpha$ radiation the density of points is about ten times as great as for Cu $K\alpha$ radiation, with the same unit cell. Moreover this effect always increases the observed intensity, in agreement with our discrepancies. Because of the geometrical factors involved, the effect, if observed for one axis of rotation, generally vanishes for an axis not equivalent crystallographically to the first. Finally, the effect involves the cooperation of two strong reflections which add up vectorially in reciprocal space to the enhanced vector. In our crystal the 200 reflection is extraordinarily intense, because of the layering. Therefore hkl reflections can be enhanced by cooperation of 200 and $h \pm 2, k, l$. For pictures made around **a**, and reflections with small k and l, many hkl and $h \pm 2, k, l$ can reach the sphere of reflection at about the same time. In view of this uncertainty it was decided to give these reflections, relatively few in number, zero weight in the refinement of parameters. On the half-cell projection the calculated values were used in place of the observed values. Cochran (1948) has shown that this process in Fourier calculations is equivalent to giving these reflections zero weight.

At this stage of the refinement, which was completed in early 1950, the discrepancy index $100 \Sigma |F_o -$
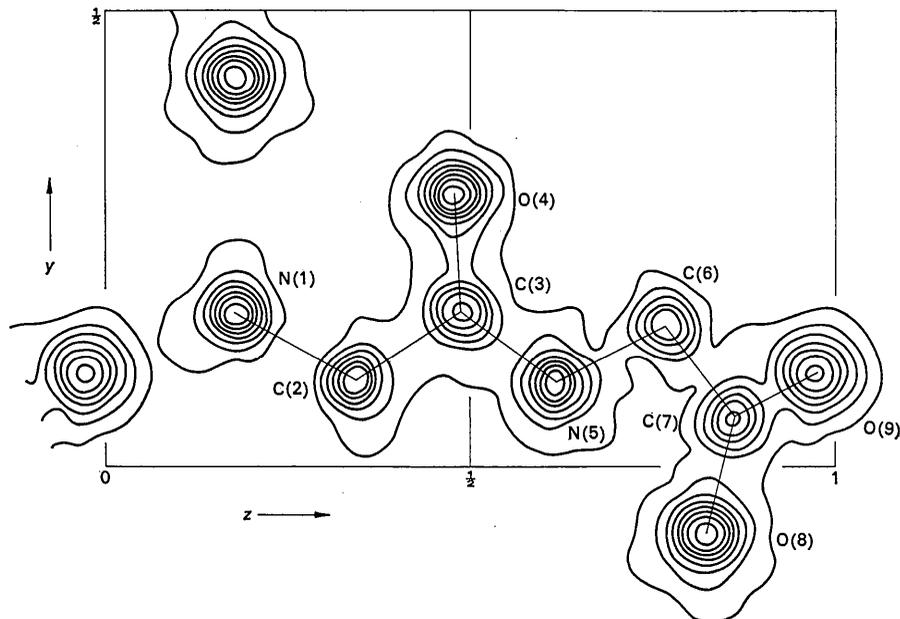
Fig. 3. Half-cell electron density projection onto (100). Arbitrary contour interval.

$F_c|/\Sigma\ |F_{obs}|$, was calculated by layers: $h0l$, 12·9; $0kl$, 17·8; $1kl$, 17·5; $3kl$, 18·6; $5kl$, 18·7; $7kl$, 22·0; and $9kl$, 25·9. The overall value was 18·1.

Refinement
with use of individual anisotropic temperature factors

In view of the more recent availability of a high speed digital computer, one of us (BDS) undertook the final refinement of the structure. The earlier least-squares parameters were adopted for further three-dimensional block-diagonal least-squares adjustment with individual anisotropic temperature factors. The latter were expressed in the form:

$$T_i = \exp[-(\alpha_i h^2 + \beta_i k^2 + \gamma_i l^2 + \delta_i hk + \epsilon_i hl + \eta_i kl)] .$$

The overall temperature factors $A_{\parallel(100)} = 1\cdot25$ Å^2 and $(A+B)_{\perp(100)} = 2\cdot13$ Å^2 were transformed by the method outlined by Rollett & Davies (1955) to correspond to the above expression.

The positional parameters of hydrogen atoms were those deduced from the half-cell structure, as listed in Table 1. An isotropic temperature factor with $B = 2\cdot0$ Å^2 was assigned to all the hydrogen atoms. Neither the positional parameters nor the temperature factor constant for hydrogen atoms were allowed to vary throughout the subsequent least-squares cycles, though their contributions to the structure factors were always incorporated. The R value for these initial parameters was 18·2. The form factor curves used were averages of Berghuis, Haanappel, Potters, Loopstra, MacGillavry & Veenendaal (1955) and Hoerni & Ibers (1954) for carbon, nitrogen, and oxygen, whereas the McWeeny (1951) form factors were used for the hydrogen atoms.

The function minimized is $\Sigma w(F_{obs}^2 - F_{calc}^2)^2$. The program allows a degree of freedom in choosing the weighting scheme and the first weighting scheme used was:

$$\sqrt{w} = 1/F_{obs} \qquad \text{when } F_{obs} > 4F_{obs\ min}$$
$$\sqrt{w} = 1/\sqrt{4F_{obs\ min} \cdot F_{obs}} \quad \text{when } F_{obs} \le 4F_{obs\ min} .$$

Six cycles of least-squares refinement were carried out with 1098 reflections (excluding the unobserved reflections and those reflections which were possibly affected by the Renninger effect) and the above weighting scheme. The R value dropped to 13·6. The shifts indicated were about one-tenth of the indicated standard deviations.

At this stage the weighting scheme was changed to:

$$\sqrt{w} = 1/F_{obs}^2 \qquad \text{when } F_{obs} > 4F_{obs\ min}$$
$$\sqrt{w} = 1/4F_{obs\ min} \cdot F_{obs} \quad \text{when } F_{obs} \le 4F_{obs\ min} .$$

This change produced shifts significant in comparison with standard deviations. Eight more cycles brought R to 12·4 for all the observed reflections and the last shifts ranged from 1/3 to 1/10 of the standard deviations. Inspection of the structure factors showed no change in sign except for a few of the unobserved reflections. However, reflection 004 was found to calcu-

late invariably less than the value of $F_{obs\ min}$. Checking of the records revealed that this reflection is indeed an unobserved reflection and the F_{obs} value taken was a clerical error. At the same time it became evident that some of the $h0l$ reflections were not on the same scale as others. The data tape was corrected to comply with these observations and the next cycle showed shifts for x and z for some atoms once again comparable to the standard deviations. Three more cycles (the 15th, 16th, and 17th) brought R to 12·3. The final shifts in positional parameters and temperature parameters were all less than one-tenth of the indicated standard deviations.

The final positional parameters for carbon, nitrogen and oxygen atoms are presented in Table 2 along with their standard deviations. The corresponding temperature factor constants are listed in Table 3. These parameters and those assigned to hydrogen atoms were used to calculate the final set of structure factors which are presented in Table 4 along with $(F_{hkl})_{obs}$.

Table 2. *Final positional parameters and their standard deviations*

	x	$10^4\sigma_x$	y	$10^4\sigma_y$	z	$10^4\sigma_z$
N(1)	0·0988$_3$	6$_5$	0·1689$_1$	3$_8$	0·1767$_5$	4$_4$
C(2)	0·2396$_9$	7$_7$	0·0947$_6$	4$_4$	0·3439$_5$	5$_3$
C(3)	0·2392$_0$	7$_7$	0·1700$_9$	4$_1$	0·4858$_8$	5$_3$
O(4)	0·2120$_1$	6$_7$	0·2992$_9$	3$_7$	0·4785$_2$	4$_8$
N(5)	0·2788$_0$	6$_8$	0·0928$_4$	3$_3$	0·6174$_3$	4$_5$
C(6)	0·3007$_8$	8$_7$	0·1553$_4$	4$_6$	0·7675$_0$	5$_7$
C(7)	0·2507$_2$	6$_3$	0·0535$_4$	3$_9$	0·8622$_2$	4$_5$
O(8)	0·2316$_2$	6$_0$	−0·0737$_9$	3$_4$	0·8213$_6$	4$_5$
O(9)	0·2362$_5$	7$_5$	0·1051$_6$	4$_5$	0·9749$_9$	5$_1$

Table 3. *Final values of the temperature-factor parameters and their standard deviations (in parentheses)*

The temperature factor for each atom is of the form

$$T_i = \exp[-(\alpha_i h^2 + \beta_i k^2 + \gamma_i l^2 + \delta_i hk + \epsilon_i hl + \eta_i kl)]$$

	$10^4\alpha$	$10^4\beta$	$10^4\gamma$	$10^4\delta$	$10^4\epsilon$	$10^4\eta$
N(1)	143	46	57	0·9	108	−0·4
	(10)	(3)	(5)	(10)	(11)	(6)
C(2)	154	45	69	33	134	11
	(13)	(4)	(6)	(10)	(14)	(7)
C(3)	134	36	70	−23	108	6
	(11)	(3)	(6)	(10)	(13)	(6)
O(4)	234	43	109	−34	206	14
	(13)	(3)	(6)	(10)	(15)	(7)
N(5)	167	32	64	−5	135	6
	(11)	(3)	(5)	(8)	(12)	(6)
C(6)	186	50	77	−26	169	−16
	(14)	(4)	(6)	(12)	(16)	(8)
C(7)	84	40	45	29	66	11
	(9)	(3)	(5)	(9)	(10)	(6)
O(8)	181	39	109	−3	191	2
	(11)	(3)	(6)	(8)	(13)	(6)
O(9)	252	85	116	−25	277	−42
	(14)	(5)	(6)	(12)	(16)	(8)

Table 4. *Observed and calculated structure factors*

The three columns in each group contain the values, reading from left to right, of *l*, F_{obs}, F_c. Reflections indicated by an asterisk were given zero weight in the least-squares calculations for reasons cited in the text.

[Table of observed and calculated structure factors — a dense multi-column numerical table organized by reflection index groups (0,0,l; 2,0,l; 4,0,l; 6,0,l; 8,0,l; 10,0,l; 12,0,l; 14,0,l; 0,1,l; 1,1,l; ... ; 7,15,l). Each group lists l, F_{obs}, and F_c values.]

Table 4 (cont.)

[Extensive structure-factor table of observed and calculated values, arranged in columns by index groups (9,13,l; 0,14,l; 1,14,l; 3,14,l; 7,14,l; 5,14,l; 2,14,l; 9,14,l; 0,15,l; 7,14,l; 1,15,l; 5,15,l; 3,15,l; 9,1,l; 0,16,l; 1,16,l; 7,1,l; 3,15,l; 9,1,l; 5,16,l; 7,16,l; 3,15,l; 1,17,l; 5,17,l; 5,17,l; 0,17,l; 1,18,l; 3,18,l; 0,18,l). Numerical values not reproduced due to illegibility.]

Discussion of the results

Nature of the packing

Because of the layer-like nature of the structure its main features are easily seen by inspection of the electron density diagrams (Figs. 1, 2, and 3). Fig. 4 is a diagram of an entire unit cell looking along **a**. Hydrogen bonds are shown by broken lines. One hydrogen bond from the peptide nitrogen atom N(5) and two hydrogen bonds from the terminal nitrogen atom N(1) are used to join molecules in the same layer. The third hydrogen bond from the terminal nitrogen atom is an interaction across layers. In addition there are the usual van der Waals interactions. Oxygen atoms O(4) and O(9) are acceptors of only one hydrogen bond each, both being in the same layer. Oxygen atom O(8) is an acceptor of two hydrogen bonds, one being within the layer and the other being from the adjacent layer. The pertinent values of hydrogen bond lengths, and angles at the donor, are listed in Table 5. The shorter van der Waals contacts are presented in Table 6.

Table 5. Hydrogen-bond distances and angles

Bonds		Acceptor at
N(1)$\cdots$O(8)	2·712 Å	$\bar{x}\bar{y}(1-z)$
N(1)$\cdots$O(9)	2·729	$xy(-1+z)$
N(1)$\cdots$O(8)	2·782	$\frac{1}{2}-x, \frac{1}{2}+y, 1-z$
N(5)$\cdots$O(4)	2·966	$\frac{1}{2}-x, -\frac{1}{2}+y, 1-z$

Angles	
C(2)–N(1)$\cdots$O(8)	87·3°
C(2)–N(1)$\cdots$O(9)	107·5
C(2)–N(1)$\cdots$O(8)	110·9
C(3)–N(5)$\cdots$O(4)	105·8
C(6)–N(5)$\cdots$O(4)	132·6

Table 6. Van der Waals contacts

The first atom of each pair is in the asymmetric unit.

C(2)$\cdots$C(7)	3·422 Å	O(4)$\cdots$O(8)	3·658 Å
C(2)$\cdots$O(8)	3·035	C(7)$\cdots$N(1)	3·284
C(2)$\cdots$O(9)	3·695	O(9)$\cdots$N(1)	3·375
C(3)$\cdots$O(8)	3·239	C(2)$\cdots$O(4)	3·258
O(4)$\cdots$C(2)	3·203	C(3)$\cdots$O(4)	3·692
O(4)$\cdots$O(8)	3·338		

Intramolecular bond distances and angles

The intramolecular bond distances and angles involving the heavy atoms are listed in Table 7, along with those of β-glycylglycine, and are shown in Fig. 5. The standard errors in the positional parameters are roughly the same, about 0·005 Å, for all the atoms. The corresponding errors are 0·007 Å in the apparent bond distances and 0·4° in the bond angles. As usual, these are to be regarded only as lower limits on the actual errors.

As far as bond lengths and angles are concerned, the molecule of the α form differs little from that in the β form. The positions of the long and short C–O bonds in the carboxyl groups are reversed between the two forms and this change is in the direction to be expected by the differences in hydrogen bonds accepted at the oxygen atoms. In the β form the long C–O(9) oxygen accepts two short (strong) hydrogen bonds and the short C–O(8) oxygen accepts one longer (weaker) hydrogen bond. In the α form there are two short bonds to one oxygen atom, O(8), and only one short bond to the other, O(9). Accordingly one expects less difference in the α form and this is found. However, the difference (0·023 Å) is significant at about the 1% level (Cruickshank & Robertson, 1953). The bond angles in the carboxyl group of the β form are completely in accord with the notion that the longer C–O(9) has less double-bond character and more negative formal charge, thus stabilizing its two hydrogen bonds at the expense of resonance energy. In the α form these angles are not so completely in accord with such an explanation but they are significantly shifted in the expected directions away from those of the β form. The longer (weaker) hydrogen bond in each structure originates at the peptide nitrogen atom, N(5). Each is at least 0·2 Å longer than the other hydrogen bonds of the structure. In the α form the acceptor is the carbonyl oxygen of the peptide group while in the β form it is an oxygen atom of the carboxyl group. This suggestion, that the peptide nitrogen atom is incapable of forming hydrogen bonds of maximum strength, is in

agreement with a current survey of hydrogen bonding in peptides (Marsh & Donohue, 1967). This, in turn, suggests that the formation of hydrogen bonds in the stabilization of the α-helix in polypeptides and proteins is no more important than the concomitant matching together of hydrophobic groups. The C(3)–N(5) bond of the peptide link is 1·328 Å, more nearly the expected value than the 1·29 Å value in the β form. It was suspected at the time of the earlier investigation that the latter was in error.

The least-squares planes

In Table 8 are listed the pertinent data for certain planes of interest. The direction cosines of the normal to the planes are with respect to an orthogonal set of axes coinciding with a^*, b, and c. The least-squares planes were calculated by the method of Schomaker, Waser, Marsh & Bergman (1958). The entire molecule is not coplanar, as was the case with the β form (excepting the terminal nitrogen, N(1), which is 0·79 Å away from the best plane of the amide group in α-glycylglycine and 0·64 Å in β-glycylglycine). The carboxyl group atoms and their α carbon atoms all lie less than 0·006 Å from their plane. However, the peptide group shows rather unexpected deviations. Most of the atoms lie about 0·03 Å from the plane. If, however, one omits atom C(6) from the calculation, the remaining four atoms (C(2), C(3), O(4), and N(5), hereafter called the amide group) are found to be coplanar, with the possible exception of C(3) which is three standard deviations out. Atom C(6) itself falls 0·105 Å out of this plane. This is a significant bending of the peptide link. The angle between the normals of the amide group plane and the carboxyl group plane is 22·4°. The rotations around the bonds C(6)–C(7), C(6)–N(5), and N(5)–C(3) to make the molecules non-coplanar are by angles of 10·5, 24·9, and 4·8°, respectively. The distortion of the peptide group must involve some loss of stability. This must be made up by better hydrogen bonding and better van der Waals interactions arising from the better packing permitted by the distortion. Experimental densities are not available but the cell constants show that the α form is significantly more dense, by about 0·013 g.cm⁻³. Moreover, for three of the four hydrogen bonds, the lengths are shorter in the α form than in the β form. (α form: 2·712, 2·729, 2·782, and 2·966 Å compared with the β form: 2·68,

2·80, 2·81, and 3·07 kX.) It is not known for certain which of the three known forms is the thermodynamically stable one at room temperature. But at 105°C the β and γ crystals are both converted spontaneously to the α form in about twenty-four hours.
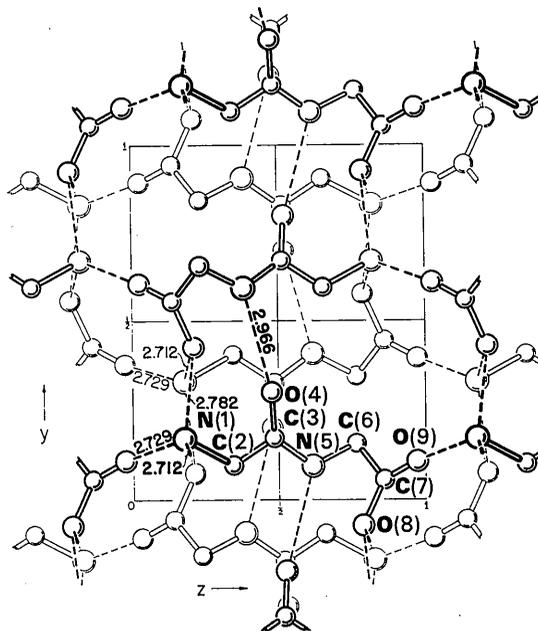
Fig. 4. Diagrammatic packing drawing. A view parallel to [100]. Broken lines indicate hydrogen bonds. Positive x is away from the viewer. Atoms in heavier lines are nearer the viewer.
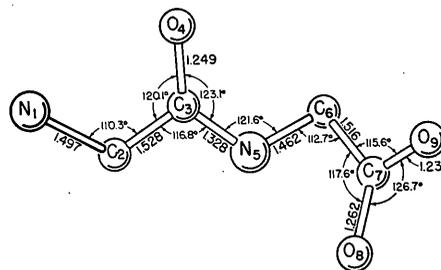
Fig. 5. Intramolecular bond lengths and angles in α-glycylglycine. N(1) is 0·79 Å above the plane of the amide group. The peptide group is significantly non-planar and the carboxyl and amide groups are inclined about 22° to each other.

Table 7. *Intramolecular bond distances and bond angles for α-glycylglycine and β-glycylglycine*
(*Hughes & Moore*, 1949)

Bond	α	β	Angle	α	β
N(1)–C(2)	1·497 Å	1·51 kX	N(1)–C(2)–C(3)	110·3°	110°
C(2)–C(3)	1·528	1·53	C(2)–C(3)–O(4)	120·1	121
C(3)–O(4)	1·249	1·23	C(2)–C(3)–N(5)	116·8	114
C(3)–N(5)	1·328	1·29	O(4)–C(3)–N(5)	123·1	125
N(5)–C(6)	1·462	1·48	C(3)–N(5)–C(6)	121·6	122
C(6)–C(7)	1·516	1·53	N(5)–C(6)–C(7)	112·7	110
C(7)–O(8)	1·262	1·21	C(6)–C(7)–O(8)	117·6	123
C(7)–O(9)	1·239	1·27	C(6)–C(7)–O(9)	115·6	112
			O(8)–C(7)–O(9)	126·7	124

The temperature factors

The temperature factor parameters for the nine heavy atoms were transformed into parameters related to the principal axes of the ellipsoids of vibration (Rollett & Davies, 1955); the magnitudes of these principal axes and their direction cosines relative to the orthogonal set coinciding with a^*, b, and c are listed in Table 9.

The vibrations implied by the temperature-factor parameters are in agreement with those to be expected on the basis of the geometry of the molecule and the intermolecular packing. The largest component of vi-

Table 8. *Direction cosines of the normals of some planes and the deviations of atoms from their best least-square plane*

Direction cosines are with respect to an orthogonal set coinciding with a^*, b, and c.

	Entire molecule Plane 1	Peptide group Plane 2	Carboxyl group Plane 3	The amide group Plane 4
cos α^*	−0·99980	−0·97895	−0·97890	−0·98228
cos β	−0·01885	−0·16121	0·14088	−0·15887
cos γ	0·00633	0·12510	−0·14798	0·09933

Atom	Distance from plane	Atom	Distance from plane	Atom	Distance from plane	Atom	Distance from plane
C(2)	0·028 Å	C(2)	−0·030 Å	C(6)	−0·002 Å	C(2)	−0·005 Å
C(3)	0·026	C(3)	0·025	C(7)	0·006	C(3)	0·017
O(4)	0·176	O(4)	0·000₆	O(8)	−0·002	O(4)	−0·005
N(5)	−0·204	N(5)	0·033	O(9)	−0·002	N(5)	−0·005
C(6)	−0·346	C(6)	−0·034	Origin	−2·538	C(6)†	−0·105
C(7)	−0·004	Origin	−1·327	N(5)†	0·247	Origin	−1·412
O(8)	0·139	N(1)†	0·608				
O(9)	0·086						
N(1)†	0·901						
Origin	−1·550						

† These atoms were given zero weight in the calculation of the plane. The rest of the atoms in each group were given weights proportional to their atomic numbers.

Table 9. *Magnitudes and direction cosines of the principal axes of the vibration ellipsoids*

	Axis	B_i	$g_i{}^{a*}$	$g_i{}^{b}$	$g_i{}^{c}$
N(1)	1	2·394 Å	0·964	0·021	−0·267
	2	1·693	−0·026	0·999₅	−0·018
	3	1·259	0·266	0·024	0·964
C(2)	1	2·656	0·922	0·375	−0·099
	2	1·505	−0·385	0·851	−0·358
	3	1·391	−0·050	0·368	0·928
C(3)	1	2·373	0·838	−0·337	−0·430
	2	1·730	0·516	0·229	0·826
	3	1·125	0·180	0·913	−0·365
O(4)	1	3·866	0·967	−0·208	−0·148
	2	2·528	0·223	0·409	0·885
	3	1·260	0·123	0·889	−0·442
N(5)	1	2·721	0·985	−0·053	−0·164
	2	1·379	0·169	0·495	0·853
	3	1·073	−0·036	0·868	−0·496
C(6)	1	3·078	0·969	−0·247	−0·011
	2	1·755	0·241	0·954	−0·181
	3	1·381	0·055	0·173	0·983
C(7)	1	1·791	0·655	0·723	−0·217
	2	1·186	−0·189	0·435	0·880
	3	1·007	−0·731	0·536	−0·422
O(8)	3	3·011	0·950	−0·007	0·313
	2	2·064	−0·311	0·103	0·945
	3	1·430	0·039	0·995	−0·096
O(9)	1	4·314	0·909	−0·342	0·241
	2	3·092	0·389	0·902	−0·186
	3	1·283	−0·154	0·263	0·953

bration for each atom is approximately perpendicular to the *bc* plane, which is nearly the plane of the entire molecule. The smallest vibration is for the atom C(7) which is covalently bonded to three heavy atoms. The atom O(9), which is involved only in one hydrogen bond, has larger vibrations than O(8), which is bonded by two hydrogen bonds.

Perdeuteroglycylglycine

Freeman, Paul & Sabine (1967) have just completed a neutron diffraction study of fully deuterated glycylglycine, in its α form. They started their least-squares refinement with our preliminary least-squares parameters. The deuteration has caused an increase in the length of *a* of approximately 0·12 Å. Despite this change the intramolecular bond lengths and angles found for the heavy atoms of the molecule confirm those reported here to within about the sums of the standard deviations. They also find the peptide group to be nonplanar. Their hydrogen (deuterium) bonds are substantially the same as ours with the exception of N(1)···O(8), across layers, which is approximately parallel to **a**. This length has increased from 2·712 to 2·758 Å. There are two of these bonds 'in tandem' in the *a* repeat and the total increase, $2 \times (2 \cdot 758 - 2 \cdot 712) = 0 \cdot 092$ Å, seems to be associated with the 0·12 Å increase in *a*. We leave the interpretation of these interesting deuteration effects to the authors who discovered them. We are grateful to Dr Freeman for a pre-publication report on their results.

References

BERGHUIS, J., HAANAPPEL, IJ. M., POTTERS, M., LOOPSTRA, B. O., MACGILLAVRY, C. H. & VEENENDAAL, A. L. (1955). *Acta Cryst.* **8**, 478.
BERNAL, J. D. (1931). *Z. Kristallogr.* **78**, 363.
BOOTH, A. D. (1945). *Trans. Faraday Soc.* **41**, 434.
COCHRAN, W. (1948). *Acta Cryst.* **1**, 138.
COCHRAN, W. & DYER, H. B. (1952). *Acta Cryst.* **5**, 634.
CRUICKSHANK, D. W. J. & ROBERTSON, A. P. (1953). *Acta Cryst.* **6**, 698.
DELANGE, J. J., ROBERTSON, J. M. & WOODWARD, I. (1939). *Proc. Roy. Soc.* A, **171**, 398.
DONOHUE, J. & SCHOMAKER, V. (1949). *Acta Cryst.* **2**, 344.
FREEMAN, H. C., PAUL, G. L. & SABINE, T. M. (1967). Private communication.
HOERNI, J. A. & IBERS, J. A. (1954). *Acta Cryst.* **7**, 744.
HUGHES, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1741.
HUGHES, E. W. (1949). *Acta Cryst.* **2**, 37.
HUGHES, E. W. & MOORE, W. J. (1949). *J. Amer. Chem. Soc.* **71**, 2618.
Internationale Tabellen zur Bestimmung von Kristallstrukturen (1935). Berlin: Borntraeger.
LU, C. S. (1943). *Rev. Sci. Instrum.* **14**, 331.
MARSH, R. E. & DONOHUE, J. (1967). *Advanc. Protein Chemistry.* In the press.
MCWEENY, R. (1951). *Acta Cryst.* **4**, 513.
RENNINGER, M. (1937). *Z. Physik*, **106**, 141.
ROLLETT, J. S. & DAVIES, D. R. (1955). *Acta Cryst.* **8**, 125.
SCHOMAKER, V. (1941). Private communication. Reported to American Society for X-Ray and Electron Diffraction, Gibson Island Meeting, 1941, by E. W. Hughes.
SCHOMAKER, V., WASER, J., MARSH, R. E. & BERGMAN, B. G. (1959). *Acta Cryst.* **12**, 600.
WILSON, A. J. C. (1942). *Nature Lond.* **150**, 151.

The Geometrical Basis of Crystal Chemistry. IX. Some Properties of Plane Nets

BY A. F. WELLS

Research Department, Imperial Chemical Industries Limited (Dyestuffs Division), Manchester 9, England

The solutions of the equations $\Sigma \, n\varphi_n = 6$, 4, 3⅓, and 3 for 3-, 4-, 5-, and 6-connected plane nets respectively represent the permissible proportions of polygons of various kinds. They give no information about the arrangement of the polygons relative one to another. An examination is made of the possible configurations of the 3-connected 5:7, 4:8, 3:9 and 4:7 nets, a family of 4-connected nets, and nets with alternate *p*- and *q*-connected points.

The general aim of these studies (Wells, 1954*a,b,c,d*, 1955, 1956, 1965) is to further our understanding of the reasons for the adoption of a particular crystal structure by a given element or compound. The amount of effort that is being devoted to building up the basic geometrical background to crystal chemistry is surprisingly small compared with that devoted to the actual determination of crystal structures. Consider,