[CONTRIBUTION NO. 1276 FROM THE GATES AND CRELLIN LABORATORIES OF CHEMISTRY, CALIFORNIA INSTITUTE OF TECHNOLOGY]

The Crystal Structure of β-Glycylglycine

BY EDWARD W. HUGHES AND WALTER J. MOORE[1]

Some time before the war the determination of the crystal structure of glycylglycine was undertaken as a continuation of the X-ray diffraction studies[2] which are part of a program of research upon the constitution and configuration of proteins. This investigation, the first concerned with a linear peptide, was stopped in 1941 because of the war, and a brief description of the results as of that time was published.[3] The work has now been completed and the details are reported below.

A preliminary investigation had been made by Bernal.[4] He grew crystals from mixtures of *n*-propyl alcohol and water and found three different crystal forms to which he applied the identifying prefixes α, β and γ. All three of these were observed growing side by side in the same mother liquor. From the unit cell constants given it seemed likely that the β-form, having a short needle axis period of only 4.62 Kx., might be the easiest structure to solve and it was accordingly chosen as the subject of the present research. This has proved to be the case so far as finding a trial structure was concerned but the crystals, being of an exceedingly thin needle-like habit, are unsuitable for obtaining good intensity data about any but the needle axis of the crystal. With the discovery of part cell projections[5] it has seemed that the α-form is best suited for the most precise determination of interatomic distances. Accordingly no effort has been made in the present research to achieve the accuracy possible with complete three-dimensional data. A careful, complete investigation of the α-form is now under way and the γ-form will be examined well enough to determine the configuration of the molecule.

In the present investigation crystals were grown from mixtures of *n*-propyl alcohol and water. The plate-like α-form was always obtained and it was only after repeated attempts that one crystallization yielded also a few needles of the γ-form. Finally after numerous further crystallizations one crop containing some β-crystals appeared. It is not known which of the three forms is stable at room temperature. It was shown by X-ray photographs that upon heating for twenty-four hours at 105° the β- and γ-forms are converted into the α-form. The exceedingly fine β-crystals show the {100} and {001} forms. The faces terminating the needles were not determined.

All X-ray photographs were made with CuKα rays with λ taken as 1.54 Kx. A few well-exposed oscillation pictures were made about c_0 to confirm the length of that axis. Complete sets of oscillation pictures were made with multiple films[6] about a_0 and b_0 for intensities. For the a_0 pictures a section of needle was broken off which was only a few tenths of a millimeter in length but even so its length was at least twice the thickness of the specimen and the resulting variation in the sizes of the X-ray spots complicated the estimation of the intensities. One zero layer b_0 Weissenberg picture was also prepared with multiple films for intensity measurements.

The monoclinic unit cell dimensions found from high order reflections on equators of oscillation photographs are a_0 = 17.89 Kx., b_0 = 4.62 Kx. and c_0 = 17.06 Kx. The angle β, from optical and X-ray goniometry, is 125°10′. Bernal's dimensions[4] were given only approximately and agree with these within his precision, except that we have found c_0 to be twice that given by him. With eight molecules per cell the calculated density is 1.512 g./cc., a reasonable value.

The only consistent extinctions observed are hkl when $k + l$ is odd and $h0l$ when h is odd. These indicate the space group Aa or $A2/a$. Pyro and piezo-electric tests proved negative and so the space group $A2/a$ was adopted. This choice is confirmed by the fact that a structure based upon it has been found which is reasonable and in good agreement with the observed intensities.

All intensities were determined by visual estimates. After correction for Lorentz and polarization factor the relative F's were derived from them by taking their square roots. These F's were later placed upon an absolute scale by comparison with the F's calculated from the final structure. They are recorded in Tables I and II, rounded to the nearest integer and arranged in order of decreasing spacing.

Determination of the Structure

The $h0l$ intensities were used to prepare a Patterson projection onto (010). As the intensities were then on an arbitrary scale F_{000} could not be included and consequently considerable areas of the cell show a negative Patterson function. Figure 1 shows by contours only the positive

---

[1] Present address: The Catholic University of America, Washington, D. C.

[2] Diketopiperazine, R. B. Corey, THIS JOURNAL, **60**, 1598 (1938); glycine, G. Albrecht and R. B. Corey, *ibid.*, **61**, 1087 (1939); DL-alanine, H. A. Levy and R. B. Corey, *ibid.*, **63**, 2095 (1941).

[3] E. W. Hughes and W. J. Moore, THIS JOURNAL, **64**, 2236 (1942).

[4] J. D. Bernal, *Z. Krist.*, **78**, 363 (1931).

[5] V. Schomaker, private communication (1942); A. D. Booth, *Trans. Faraday Soc.*, **41**, 434 (1945).

[6] J. J. Delange, J. M. Robertson and I. Woodward, *Proc. Roy. Soc.* (*London*), **A171**, 398 (1939).

portions in one-eighth of a cell. The calculations were carried out with Lipson–Beevers strips at intervals of one-hundred-twentieth of a cell edge.

TABLE I

STRUCTURE FACTORS OF THE $h0l$ REFLECTIONS

$h0l$	F obs.	F calcd.	$h0l$	F obs.	F calcd.	$h0l$	F obs.	F calcd.
$20\bar{2}$	5	+ 6	606	26	−30	14,0,2	5	− 6
200	29	+ 28	804	8	+ 6	$18,0,\bar{6}$	10	+ 5
002	54	− 61	$12,0,\bar{2}$	10	−11	$18,0,\overline{14}$	<4	+ 2
$40\bar{2}$	64	− 75	408	60	−61	$10,0,\overline{18}$	14	−12
$20\bar{4}$	56	− 67	10,0,2	6	+ 9	$12,0,\overline{18}$	13	+13
202	34	+ 36	$12,0,\overline{12}$	12	+11	$8,0,\overline{18}$	<3	0
$40\bar{4}$	40	− 49	$2,0,\overline{12}$	18	−25	$2,0,\overline{16}$	10	+11
400	184	−219	$14,0,\bar{8}$	7	− 3	$14,0,\overline{18}$	21	+24
004	17	− 8	$14,0,\bar{6}$	16	−10	2,0,14	14	+12
$60\bar{4}$	41	+ 44	$14,0,\overline{10}$	30	+42	$18,0,\bar{4}$	<4	− 2
$40\bar{6}$	29	+ 37	2,0,10	12	+14	16,0,0	<3	− 1
$60\bar{2}$	39	− 45	$8,0,\overline{14}$	93	+87	$6,0,\overline{18}$	5	+ 2
$20\bar{6}$	54	+ 49	12,0,0	18	−17	$18,0.\overline{16}$	13	−14
402	66	+ 63	$14,0,\bar{4}$	<5	−11	10,0,8	14	+13
204	68	− 60	$10,0,\overline{14}$	32	−37	$16,0,\overline{18}$	10	+ 9
$60\bar{6}$	4	+ 14	$6,0,\overline{14}$	26	+29	8,0,10	20	+23
600	7	+ 5	$14,0,\overline{12}$	5	− 6	$20,0,\overline{12}$	<4	0
006	76	− 60	$12,0,\overline{14}$	82	−77	12,0,6	6	− 7
$80\bar{4}$	60	+ 60	0,0.12	13	+ 7	$20,0,\overline{10}$	6	− 9
$80\bar{6}$	112	+104	806	8	−10	$20,0,\overline{14}$	8	− 8
$40\bar{8}$	32	− 17	608	14	−11	14,0,4	9	−14
$60\bar{8}$	4	+ 6	$4,0,\overline{14}$	49	−42	6,0,12	8	+ 8
404	28	− 30	10,0,4	22	+23	0,0,16	<4	0
$80\bar{2}$	40	+ 33	$14,0,\bar{2}$	33	+42	$20,0,\bar{8}$	7	+ 5
602	6	+ 6	4,0,10	13	− 6	$4,0,\overline{18}$	5	− 5
$20\bar{8}$	35	− 33	$16,0,\overline{10}$	5	+ 2	$18,0,\bar{2}$	4	− 1
$80\bar{8}$	83	− 77	$16,0,\bar{8}$	<3	+ 3	$18,0,\overline{18}$	6	− 3
206	12	+ 10	$14,0,\overline{14}$	<3	+ 2	$12,0,\overline{20}$	<5	+ 7
800	33	+ 25	12,0,2	7	− 5	4,0,14	5	+ 7
$10,0,\bar{6}$	<5	− 13	$16,0,\bar{6}$	28	+26	$20,0,\bar{6}$	22	−26
008	112	+111	$16,0,\overline{12}$	7	− 6	$14,0,\overline{20}$	<3	− 3
$10,0,\bar{4}$	35	+ 40	$2,0,\overline{14}$	<3	+ 2	$20,0,\overline{16}$	3	+ 8
$6,0,\overline{10}$	4	− 8	$10,0,\overline{16}$	5	− 5	$10,0,\overline{20}$	8	+10
$10,0,\bar{8}$	5	− 13	$8,0,\overline{16}$	33	−32	16,0,2	10	− 8
$4,0,\overline{10}$	57	− 55	2,0,12	18	−15	$8,0,\overline{20}$	14	−13
$8,0,\overline{10}$	13	+ 10	$12,0,\overline{16}$	5	− 1	$16,0,\overline{20}$	<5	+ 3
$10,0,\bar{2}$	44	− 41	14,0,0	18	−20	$2,0,\overline{18}$	5	− 6
604	18	+ 16	$16,0,\bar{4}$	31	+30	$20,0,\overline{18}$	20	−21
406	25	+ 18	$6,0,\overline{16}$	<4	− 5	$20.0,\bar{4}$	8	−16
802	17	+ 26	$16,0,\overline{14}$	21	+19	2,0,16	13	+15
$2,0,\overline{10}$	18	− 17	$14,0,\overline{16}$	8	+ 9	$22,0,\overline{12}$	<2	− 2
$10,0,\overline{10}$	33	− 32	808	33	+33	18,0,0	<2	− 2
208	47	+ 48	10,0,6	20	−20	10,0,10	3	+12
$12,0,\bar{6}$	72	− 84	0,0,14	16	+15	12,0,8	17	−24
$12,0,\bar{8}$	20	+ 13	6,0,10	8	− 8	$6,0,\overline{20}$	6	+ 5
10.0,0	9	− 3	12,0.4	<3	− 1	$18,0,\overline{20}$	12	−12
$12,0,\bar{4}$	24	− 22	$4,0,\overline{16}$	17	+16	$22,0,\overline{10}$	10	−13
$8,0,\overline{12}$	9	+ 10	$18,0,\overline{10}$	9	+ 5	$22,0,\overline{14}$	<2	+ 2
$6,0,\overline{12}$	28	+ 34	$16,0,\bar{2}$	4	+ 4	8,0,12	9	− 4
0,0,10	5	+ 4	$18,0,\bar{8}$	15	−14	14,0,6	17	+25
$12,0,\overline{10}$	24	+ 18	$18.0,\overline{12}$	11	−13	$22,0,\overline{16}$	16	+23
$10,0,\overline{12}$	22	+ 23	$16,0,\overline{16}$	15	+14	$22,0,\bar{8}$	17	+31
$4,0,\overline{12}$	<4	− 1	4,0,12	13	−16			

Normally one would hardly expect to be able to interpret a Patterson projection for an organic crystal with thirty-six parameters. In the present case, however, the shortness of the projection axis introduced considerable simplification and the interpretation was relatively easy and straightforward. For dicyandiamide,[7] which also has such a short axis, it was shown that a certain area about the origin contains only peaks originating inside single molecules. Following the same argument and calculation we find that in Fig. 1 the smaller dotted circle about the origin shows the area outside of which all hydrogen-bonded intermolecular peaks must lie while the larger dotted circle shows the area outside which all van der Waals intermolecular peaks must lie, as well as all non-contact peaks. Moreover, the peaks are distributed roughly along the lines $X = 0$ and $X = 1/4$, with generally negative regions in between along the line $X = 1/8$. Bernal[4] reported the crystal optically positive with the large refractive index obtaining when the electric vector is parallel to c_0. These facts all suggest that the long way of the molecule is parallel to c_0 and indeed one can see in the arrangement of peaks along $X = 0$ the suggestion of an extended zigzag chain molecule. Thus in addition to the restrictions regarding the dotted

TABLE II

STRUCTURE FACTORS OF THE $0kl$ REFLECTIONS

$0kl$	F obs.	F calcd.	$0kl$	F obs.	F calcd.
002	54	− 61	042	<9	− 9
011	31	+ 33	044	15	−17
004	17	− 8	039	<9	− 5
013	21	+ 16	0,1,13	27	+25
015	52	+ 52	0,2,12	19	+22
006	76	− 60	046	<9	+ 1
020	15	+ 15	0,0,14	16	+15
022	<6	− 2	0,3,11	16	+17
024	19	+ 22	048	15	−13
017	<7	+ 6	0,2,14	17	−23
008	112	+111	051	8	+ 8
026	34	− 42	0,1,15	<8	− 9
031	24	− 24	053	10	−10
019	27	+ 30	0,4,10	<8	− 5
033	<8	+ 5	0,3,13	<8	+ 3
0,0,10	5	+ 4	055	30	−31
028	<8	+ 4	0,0,16	<4	0
035	<8	− 7	057	8	+12
0,1,11	<9	− 10	0,4,12	16	−13
037	48	− 45	0,2,16	<7	+ 1
0,2,10	<9	+ 7	0,1,17	<7	+ 7
0,0,12	13	+ 7	0,3,15	22	−19
040	32	− 24	059	8	−10

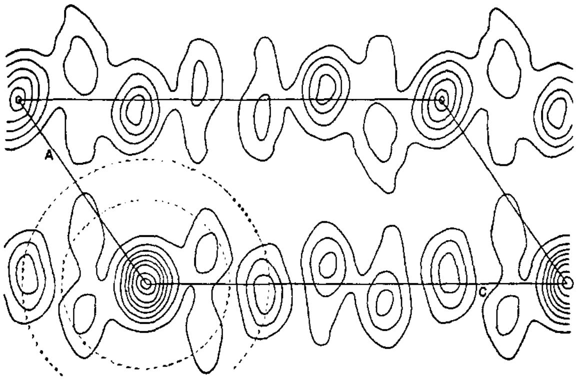


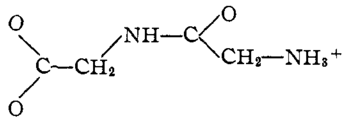
Fig. 1.—Patterson projection on (010): one-eighth of a cell from 0 to $c_0/2$ and from 0 to $a_0/4$, containing one asymmetric unit. Negative contours are omitted.

---

[7] E. W. Hughes, THIS JOURNAL, 62, 1258 (1940).

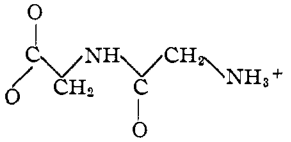
circles about the origin it appeared fairly certain that *all* the peaks along $X = 0$ arise inside single molecules, or between a molecule and the one related to it by the A centering. The size of the b_0 spacing and the optical properties also suggest that the plane of the molecule is tipped at a considerable angle to (010). We were thus led to try extended planar models with their long axes parallel to c_0 and their planes making an angle of 45° with (010). The two models tried were

I



suggested by Bernal and

II



which seemed more likely because it does not involve such a short distance between the CH_2 groups. Both models were considered to be approximately coplanar.

From the projections of these models parallel to b_0, the patterns of the expected peaks along $X = 0$ were calculated, plotted on tracing paper and superimposed on Fig. 1. Only the second model gave a good fit with the actual Patterson. The orientation of the model was adjusted to give the best possible fit with the $X = 0$ peaks and there remained only the operation of determining the X and Z parameters of the model as a whole. This was done by considering the peaks along $X = 1/4$. The array of peaks to be expected between two molecules related by a center of symmetry, which is the only symmetry element in the b_0 projection of $A2/a$, was plotted in duplicate on two sheets of tracing paper and these were then superimposed with proper orientation on Fig. 1. They were then shifted about in such a way as to maintain the centers of symmetry in the Patterson projection at (1/4,0) and (1/4, 1/4) and an effort was made to find an arrangement that would account for the peaks in the Patterson along $X = 1/4$. This did not prove difficult and in this way the approximate positions of the molecules with respect to the center of symmetry were determined. Structure factors calculated from this arrangement showed such good agreement with the observed F's that it was possible to proceed at once with the refinement of the structure by successive Fourier syntheses and the method of least squares.[8] Although in the Fourier projection five of the atoms are very well resolved there are two C–O groups which are just barely shown as separate maxima. The parameters for these were determined by analyzing the peaks as well as possible into two symmetrical peaks. The least squares refinements were made as an additional check upon these less certain parameters. The Fourier syntheses were carried out on I.B.M. Co. machines using punched cards[9] at intervals of 0.01 a_0 and 0.002 c_0. The final Fourier, made with signs agreeing with those calculated from the final parameters, is given in Fig. 2, which shows one asymmetric unit. The x and z parameters from the Fourier are recorded in Table III.

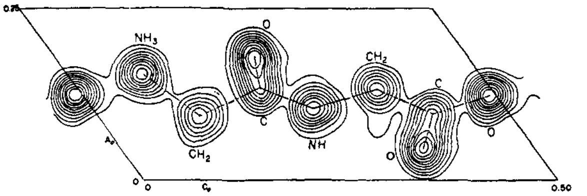


Fig. 2.—Fourier projection on (010): one-eighth of a cell containing one asymmetric unit. The contour interval is in arbitrary units.

In the least squares calculations, 100 $h0l$ reflections were used and the 100 observational equations in eighteen variables were reduced to normal equations by punched card calculations with the I.B.M. Co. machines,[9] the time required being about one day. This eighteenth order set of simultaneous linear equations was solved by an iteration method in about four hours. The least squares refinement was based on the parameters from the second from the last Fourier synthesis. The average shift was 0.0017 of a cell edge and the parameters obtained differed from those of Table III on the average by 0.0005 of a cell edge, with a maximum discrepancy of 0.002. For the four atoms not perfectly resolved in the Fourier the least squares parameters differ from the Fourier results by an average of only 0.0004 of a cell edge, with a maximum discrepancy of only 0.001 of a cell edge.

In order to obtain trial y parameters, a model of expected interatomic distances and bond angles was oriented in space so as to project onto the x, z position and the y parameters of all atoms were calculated with respect to an arbitrary zero point in the molecule. The model was then translated

TABLE III

ATOMIC COÖRDINATES

Atom	x	y	z
O-9	0.124	0.375	0.490
O-8	.047	.670	.367
O-4	.179	.070	.241
C-3	.132	.290	.219
C-7	.095	.460	.406
CH_2-6	.134	.280	.362
CH_2-2	.093	.440	.123
NH-5	.105	.400	.268
NH_3-1	.155	.390	.091

---

[8] E. W. Hughes, THIS JOURNAL, **63**, 1737 (1941), particularly p. 1741.

[9] P. A. Schaffer, Jr., V. Schomaker and L. Pauling, *J. Chem. Phys.*, **14**, 648 (1946).

TABLE IV

INTERATOMIC DISTANCES

Chemical bonds		van der Waals	
$NH_3$1–$CH_2$2	1.51 Kx.	O9A......$CH_2$2	3.32 Kx.
$CH_2$2–C3	1.53	O4B......$CH_2$2	3.39
C3–O4	1.23	$CH_2$6......O9C	3.57
C3–NH5	1.29	$CH_2$2......O8D	3.17
NH5–$CH_2$6	1.48	$CH_2$2......O9D	3.50
$CH_2$6–C7	1.53	$CH_2$2*b*.....O4	3.37
C7–O9	1.27	$CH_2$6*b*.....O8	3.25
C7–O8	1.21	$CH_2$6......$CH_2$6C	4.11
Hydrogen bonds		$CH_2$2......$CH_2$2E	3.59
$NH_3$1...O9A	2.68 Kx.	O8D......C3	3.63
$NH_3$1...O9Ab[a]	2.80	C3........O4B	3.31
$NH_3$1...O4B	2.81	C3........O8D*b*	3.40
NH5....O8D	3.07	C3b.......O4	3.67
		C7........O8	3.72
		$CH_2$6......$NH_3$1B	3.83
		$NH_3$1.....$CH_2$2E	3.87
		NH5......NH5D*b*	3.71
		O9*b*.......O8	3.69

[a] A *b* suffix indicates an atom one b_0 repeat above or below the atom indicated without the suffix.

as a whole along b_0 until reasonable intermolecular distances were obtained and rough calculations of some 0*kl* intensities were in qualitative agreement with Table II.

The *y* parameters so obtained were refined by two successive least squares refinements using all the 0*kl* data. Because of the shortness of b_0, these parameters have been given only to the nearest 0.005, since this corresponds to about the same absolute accuracy in Kx. units as for the *x*, *z* coördinates. The final values are recorded in Table III.

As a final check upon the parameters, they were used to calculate values of the *F*'s for comparison with the experimental *F*'s. The scattering factors used were taken from the "Internationale Tabellen zur Bestimmung von Kristallstrukturen" with the following modification made in an effort to allow in part for the hydrogens. For $>CH_2$ we used the scattering power of C plus the difference between the scattering powers of O and $O^=$. A similar correction was made to N to get $-NH_3^+$. For $>NH$ half this correction was applied to N. The temperature factor had $B = 2$. These scattering factors were also used in the least squares calculations. The corrections used here were devised with the idea in mind that the extra electrons brought to each atom by the hydrogen attached to it are distributed in the outer portions of the atom in much the same way as are the extra electrons in an oxygen ion and will therefore modify its scattering power in the same general way. It is to be noted that while these extra electrons *increase* the scattering power at small scattering angles they actually *decrease* it at moderate scattering angles. It is felt that this sort of correction, although obviously very approximate, is much better than either ignoring the hydrogens, or making a blanket percentage increase in the scattering power at all scattering angles. These latter are two procedures commonly followed in the past by others. Our correction produces on the average an improvement in agreement at small and moderate scattering angles although for some reflections it may make the agreement worse because the atoms are not spherically symmetrical and the extra electrons added are not disposed symmetrically about the atoms. The ideal solution would be to include the hydrogens directly in the calculation at their proper location.

For the *h*0*l* zone the ratio $\Sigma|F_{\text{calcd}} - F_{\text{obs}}|/\Sigma|F_{\text{obs}}|$ is 0.16 and for the 0*kl* zone the value is 0.12, in each case omitting non-observably small *F*'s. It is hard to say why, as judged by this criterion, the agreement appears to be better for the zone for which the intensities ought to be less reliable. It is in part due to the fact that the exposure was much heavier in relation to the average size of the *F*'s for the *h*0*l* zone and consequently there is a relatively large number of small *F*'s actually observed. The percentage errors are greater for these small *F*'s. Upon omitting the smaller *F*'s for *h*0*l* up to a relative size comparable with the lower limit observed for 0*kl*, the ratio falls from 0.16 to 0.13. It thus appears that for this reason this ratio is not a wholly satisfactory method for judging reliability. In any event these ratios are comparable with the better values reported by other investigators for organic crystals. It is interesting to note that if we average the *F*'s for each zone, counting unobservably small *F*'s at half their maximum possible value, the ratio $\bar{F}_{h0l}/\bar{F}_{0kl} = 1.46 = \sqrt{2.1}$, which agrees with $\sqrt{2}$, the predicted[10] value, within experimental error.

A statistical study of the calculated and observed *F*'s, considering the former to represent "most probable" values, indicates a probable error of approximately ±3.0 electrons in an observed *F* of unit weight. From this and the normal equations of the least squares refinements one can derive probable errors in the coördinates of the atoms. These vary from about ±0.014 Kx. to ±0.024 Kx. for *y*'s and ±0.009 Kx. to ±0.014 Kx. for *x*'s and *z*'s. Applying these to the various bonds leads to probable errors of from ±0.015 Kx. to ±0.02 Kx. depending upon the orientation of the bond and the identity of the atoms. We thus see that for distances which are completely independent, not more than one in twenty will be in error by more than ±0.06 Kx., and half will have an error of less than ±0.02 Kx.

Discussion

The molecule is shown in Fig. 3 and the more important distances are set forth in Table IV. The numbering system for indicating atoms is shown in Fig. 4, which is a normal projection of a

---

(10) E. W. Hughes, *Acta Cryst.*, **2**, in press (1949).

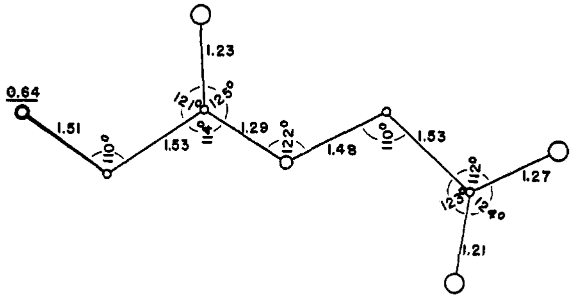


Fig. 3.—The β-glycylglycine molecule: small circles represent carbons, intermediate circles, nitrogens and large circles, oxygens. The molecule is coplanar except for the terminal nitrogen which is 0.64 Kx. above the plane.

portion of the structure onto (010). All these distances are comparable with those previously found in other structures with the following few exceptions.

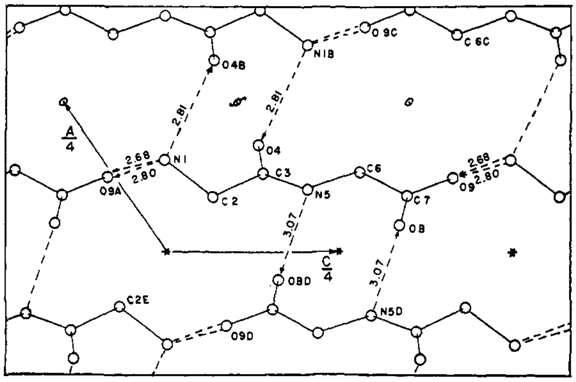


Fig. 4.—A portion of the β-glycylglycine structure projected onto (010): the figure shows the atom numbering system used in Table IV. Dotted lines indicate hydrogen bonds and associated numbers show their lengths.

The C–N$^+$ distances of 1.48 and 1.51 Kx. are definitely longer than the corresponding distances reported by Corey and co-workers,[1] who found distances from 1.39 to 1.42 Kx. Consequently, in this work every effort was made to test the possibility of the shorter distances. They were put into the trial models originally and at several stages in the refinement were re-introduced by moving one or more atoms in various ways. Subsequent refinement invariably raised these C–N distances to those quoted.

They are comparable with 1.465 Kx. in methylamine hydrochloride,[11] with 1.49 Kx. in geranylamine hydrochloride,[12] and with 1.47, 1.48 and 1.51 Kx. occurring in potassium penicillin.[13] A recent reinvestigation of DL-alanine in these laboratories by three-dimensional Fourier synthesis has revised the C–N value for this crystal[14] to 1.49 Kx. It thus appears that in amino acids, simple peptides and related substances the observed C–N$^+$ bond falls in the range from 1.46 to about 1.51 Kx. and does not seriously differ from the bond length table value of 1.47 Kx.

The other C–N bond, 1.29 Kx., which involves single bond–double bond resonance might be expected from other structures[15] to be about 1.35 Kx. and the present value may represent an error.

The difference between the two C–O distances in the carboxyl group, 1.21 and 1.27 Kx., may be outside experimental error and if so is to be explained by the fact that the longer distance is to an oxygen, O9, which has two hydrogen bonds directed to it while the shorter is to an oxygen, O2, having only one hydrogen bond directed to it (see Figs. 4 and 5 for the disposition of hydrogen bonds). This arrangement of hydrogen bonds will be stabilized if O9 has a more negative formal charge than does O2. This in turn represents a stabilization of the double bond in the C7-O2 position with a resulting shortening of the C7-O2 bond and lengthening of the C7-O9 bond as compared to the ordinary carboxyl C–O distance.

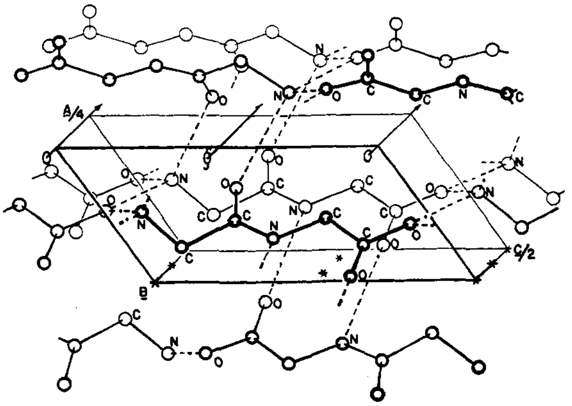


Fig. 5.—A portion of the β-glycylcine structure in oblique projection: viewed at an angle of 15° from the b_0 axis. One eighth of a cell is outlined and it contains one molecule. The remaining molecules shown are outside the part-cell and show the nature of the packing about the first molecule. For the lower two-thirds of the figure the molecules are viewed at about 45° to their own plane; in the upper third at about 25°. Dotted lines indicate hydrogen bonds.

The bond angles, shown on Fig. 3, are in agreement with predicted values. Those at the CH$_2$ atoms are tetrahedral within experimental error. The O4-C3-N5 angle of 125° and O9-C7-O2 angle of 124° are very close to the predicted value of 125°28′. The other two angles at each C-atom are expected to be 117°. For the carboxyl C-atom the values of 123 and 112° diverge from 117° in a way consistent with the notion that the

---

[11] E. W. Hughes and W. N. Lipscomb, THIS JOURNAL, **68**, 1970 (1946).

[12] G. A. Jeffrey, *Proc. Roy. Soc.* (*London*), **A183**, 388 (1945).

[13] D. Crowfoot and B. W. Rogers-Low, "The Chemistry of Penicillin," Princeton University Press, Princeton, N. J., 1949, p. 346.

[14] J. Donohue, THIS JOURNAL, **71**, in press (1949).

[15] See summary in ref. 8, p. 1745.

resonating double bond tends to be stabilized in the C7-O2 position as discussed above. At the carbonyl carbon the angles 121 and 114°, are displaced from 117° in an entirely analogous way, although to a lesser extent. This would suggest that the resonating double bond here is stabilized somewhat in the C3-O4 position as compared with the C3-N5 position. Although the C3-O4 bond is slightly shorter than the values often found for carbonyl C–O, the C3-N5 bond is also too short, as noted above, although not more than the limit of error for this determination.

The molecule is coplanar to within about ±0.03 Kx., which is within experimental error, except for the terminal NH_3 group, which falls out of the plane by 0.64 Kx. The normal to the plane makes an angle of 53°34′, with b_0 and 91°25′ with c_0.

The disposition of hydrogen bonds, shown in Figs. 4 and 5, supports the view that in β-glycylglycine the molecule is a zwitterion. The terminal nitrogen atom is bound to three neighboring oxygen atoms by hydrogen bonds and the angles formed by these bonds with each other and with the $^+NH_3$–CH_2– bond are approximately tetrahedral. This in turn strongly suggests that the N-atom carries three hydrogens tetrahedrally arranged, as one would expect for a zwitterion. Moreover, if the labile hydrogen were attached to one of the carboxyl oxygens one would expect a much greater difference between the two C–O distances than that observed.

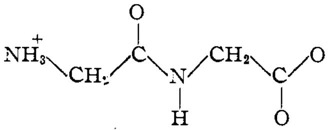
The molecules are bound together in all three directions by hydrogen bonds. Molecules are joined together in pairs by two such bonds related to each other by one of the centers of symmetry. These pairs are bound together in the a_0 and b_0 directions by spirals of hydrogen bonds about the twofold screw axes and are further bound together in the b_0 and c_0 directions by zig-zag lines of hydrogen bonds along b_0 involving a carboxyl oxygen of one molecule and the NH_3^+ atom of a molecule related to the first by *A* centering. These relationships are shown in Fig. 5 which is an oblique projection showing how the structure appears when viewed at an angle of 15° from the b_0 axis. The remaining shorter contacts between molecules are at the usual van der Waals distances, as shown in Table IV. Of these the only ones which are at the lower limits observed in other crystals are the O...CH_2 group of 3.17 and 3.25 Kx. and the CH_2...CH_2 of 3.59 Kx. It thus appears that the packing is dominated by the formation of hydrogen bonds, of which the molecules form the maximum number possible.

It is perhaps of some significance in speculations regarding the possible configurations of polypeptide chains in proteins to note that this simplest of all possible peptides is able to find three different arrangements which are so nearly equal in stability that no interconversion seems to take place at room temperature. It will be interesting to discover through the X-ray investigations of α- and γ-glycylglycine now underway if these three crystal modifications involve any significant variation in the configuration of the glycylglycine molecule itself or if they merely represent three different modes of packing together molecules of the same configuration.

Summary

The crystal structure of the β-modification of glycylglycine has been investigated by the methods of X-ray diffraction. The monoclinic unit has a_0 = 17.89 Kx., b_0 = 4.62 Kx., c_0 = 17.06 Kx., β = 125°10′. The space group is $A2/a$ and there are eight molecules per cell.

The atomic parameters were refined by Fourier syntheses and least squares methods. The molecule found has the configuration



$$\overset{+}{NH_3}\diagdown CH_2\diagup\overset{\overset{\displaystyle O}{|}}{C}\diagdown\underset{\underset{\displaystyle H}{|}}{N}\diagup CH_2\diagdown\underset{\underset{\displaystyle O}{|}}{C}\diagup O$$

which is coplanar within experimental error except for the NH_3^+ group, which lies 0.64 Kx. out of the plane of the other atoms. The zwitterion form is established by the disposition of hydrogen bonds. The latter are also arranged in such a way as to bind the crystal together in all directions and they may be said to dictate the mode of packing of the molecules. The short C–N$^+$ distance previously reported in amino acids is not substantiated by this work; the interatomic distances all appear to be normal within the limits of possible error.

PASADENA 4, CALIFORNIA RECEIVED MARCH 3, 1949